

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:33:41 ; Search time 27 Seconds
(without alignments)
2192.437 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTVRLKLYSLRWISDHELYL.....AQHQHYTHMGHFIKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents AA:*

- 1: /cgn2_6/ptodata/1/iaa/5_COMB.pep.*
- 2: /cgn2_6/ptodata/1/iaa/6_COMB.pep.*
- 3: /cgn2_6/ptodata/1/iaa/H_COMB.pep.*
- 4: /cgn2_6/ptodata/1/iaa/PTUS_COMB.pep.*
- 5: /cgn2_6/ptodata/1/iaa/RE_COMB.pep.*
- 6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	755	4	PCT-US93-07923-3
2	3877	100.0	759	4	PCT-US93-07923-2
3	3877	100.0	766	1	US-08-230-491A-3
4	3877	100.0	766	1	US-08-619-280A-3
5	3877	100.0	766	1	US-08-940-391-3
6	3877	100.0	766	2	US-09-794-236-1
7	3877	100.0	766	2	US-09-265-606-3
8	3877	100.0	775	2	US-09-949-016-10450
9	3871	99.8	766	2	US-10-002-593-6
10	3871	99.8	766	2	US-09-949-016-6146
11	3871	99.8	766	2	US-10-423-714-6
12	3867	99.7	766	2	US-09-518-550-27
13	2948	76.0	593	4	PCT-US93-07923-11
14	2138	55.1	760	1	US-08-230-491A-2
15	2138	55.1	760	1	US-08-619-280A-2
16	2138	55.1	760	1	US-08-940-391-2
17	2138	55.1	760	2	US-09-265-606-2
18	2136.5	55.1	761	2	US-09-518-550-26
19	1229.5	31.7	547	2	US-09-949-016-8330
20	1217	31.4	786	2	US-09-976-674-5
21	1198	30.9	743	2	US-10-363-937-4
22	1196	30.8	706	2	US-09-976-674-41
23	1152.5	29.7	691	2	US-09-976-674-43
24	987	25.5	771	2	US-09-079-592-2
25	963.5	24.9	771	2	US-09-462-845-2
26	908	23.4	818	2	US-09-976-674-23
27	908	23.4	818	2	US-10-402-312-3

28	908	23.4	818	2	US-10-401-437-3	Sequence 3, Appli
29	908	23.4	818	2	US-10-402-067-3	Sequence 3, Appli
30	908	23.4	818	2	US-10-401-436-3	Sequence 3, Appli
31	809	20.9	676	2	US-09-518-550-42	Sequence 42, Appli
32	809	20.9	723	2	US-09-518-550-29	Sequence 29, Appli
33	791.5	20.4	710	2	US-09-518-550-28	Sequence 28, Appli
34	726	18.7	931	2	US-09-079-592-11	Sequence 11, Appli
35	591	15.2	323	2	US-09-270-767-45296	Sequence 45296, A
36	584.5	15.1	494	2	US-09-248-796A-19472	Sequence 19472, A
37	525	13.5	882	2	US-09-976-674-1	Sequence 1, Appli
38	525	13.5	882	2	US-10-070-464-1	Sequence 1, Appli
39	493.5	12.7	526	2	US-09-248-796A-19924	Sequence 19924, A
40	493	12.7	732	2	US-09-518-550-30	Sequence 30, Appli
41	488	12.6	879	2	US-09-976-674-33	Sequence 33, Appli
42	488	12.6	879	2	US-09-976-674-35	Sequence 35, Appli
43	481.5	12.4	863	2	US-09-976-674-3	Sequence 3, Appli
44	481.5	12.4	892	2	US-09-976-674-23	Sequence 23, Appli
45	481.5	12.4	892	2	US-09-976-674-27	Sequence 27, Appli

ALIGNMENTS

RESULT 1
PCT-US93-07923-3
; Sequence 3, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 502 or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janis K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 755
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
PCT-US93-07923-3

Query Match 100.0%; Score 3877; DB 4; Length 755;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDFGHSINDYSISPD 60
Db 40 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDFGHSINDYSISPD 99
QY 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 100 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 159
QY 121 DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEVFSAYSAWMSPNGTFLAYAQFND 180
Db 160 DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEVFSAYSAWMSPNGTFLAYAQFND 219
QY 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLVSTNATSIQITAP 240
Db 220 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLVSTNATSIQITAP 279
QY 241 ASMLIGDHYLDCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 300
Db 280 ASMLIGDHYLDCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 339
QY 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 340 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 399
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 400 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 459
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFTILNETKFWQMIL 480
Db 460 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFTILNETKFWQMIL 519
QY 481 PPHFDKSKKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDGRSGYQGDKI 540
Db 520 PPHFDKSKKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDGRSGYQGDKI 579
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
Db 580 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSMLVSGSGGVFKCG 639
QY 601 IAVAPVSRWEYSDVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 640 IAVAPVSRWEYSDVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 699
QY 661 VHFQSAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIKQCFSLP 716
Db 700 VHFQSAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIKQCFSLP 755
```

RESULT 2

```
PCT-US93-07923-2
; Sequence 2, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; OPERATING SYSTEM: IBM PS/2 Model 502 or 55SX
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
```

```
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janle K.
; REGISTRATION NUMBER: 34,819
; REFERENCE/DOCKET NUMBER: 00530/055002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 542-5070
; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 759
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; PCT-US93-07923-2

Query Match 100.0%; Score 3877; DB 4; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDFGHSINDYSISPD 60
Db 44 NTYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDFGHSINDYSISPD 103
QY 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 104 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 163
QY 121 DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEVFSAYSAWMSPNGTFLAYAQFND 180
Db 164 DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVEVFSAYSAWMSPNGTFLAYAQFND 223
QY 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLVSTNATSIQITAP 240
Db 224 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSSLVSTNATSIQITAP 283
QY 241 ASMLIGDHYLDCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 300
Db 284 ASMLIGDHYLDCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCILVARQHIE MST 343
QY 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 344 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 403
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 404 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 463
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFTILNETKFWQMIL 480
Db 464 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFTILNETKFWQMIL 523
QY 481 PPHFDKSKKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDGRSGYQGDKI 540
Db 524 PPHFDKSKKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDGRSGYQGDKI 583
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
Db 584 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSMLVSGSGGVFKCG 643
QY 601 IAVAPVSRWEYSDVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 644 IAVAPVSRWEYSDVYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 703
QY 661 VHFQSAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIKQCFSLP 716
Db 704 VHFQSAQISKALVDVGVDFQAMWYTTDEDHGIIASSTAHOIYTHMSHFIKQCFSLP 759
```

```
RESULT 3
US-08-230-491A-3
; Sequence 3, Application US/08230491A
; Patent No. 5587299
; GENERAL INFORMATION:
; APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
; APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
; TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN AND USES
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FELFE & LYNCH
; STREET: 805 THIRD AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: WORDPERFECT - ASC II
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/230,491A
; FILING DATE: 20-APRIL-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5587299man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 330
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-230-491A-3
Query Match 100.0%; Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKXLSRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
Db 51 NTYRLKXLSRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
Qy 61 GQFTLLVYVYKQWRHSYTSYDIYDLNKRQLITEERIPIINNTQWVWSPVGHKLAYVWNN 120
Db 111 GQFILLEVYVYKQWRHSYTSYDIYDLNKRQLITEERIPIINNTQWVWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSVRIWTGKEDIYNGITDWTVEEYVFSAYSLWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSVRIWTGKEDIYNGITDWTVEEYVFSAYSLWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDSLQPKTVRPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQPKTVRPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGHYLCVDTWATQERISLOWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGHYLCVDTWATQERISLOWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPFTLDGNSFYKIIISNEEGRHICVFOIDKDKCTFITKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPFTLDGNSFYKIIISNEEGRHICVFOIDKDKCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYLYISNEYKGMPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 51 NTYRLKXLSRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110

Db 411 TSDLYLYISNEYKGMPCGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCSPGLPLATYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDFIILNETKMYQML 480
Db 471 RCSPGLPLATYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDFIILNETKMYQML 530
Qy 481 PPFPDKSKYPLLLDDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDGRSGYQSGDKI 540
Db 531 PPFPDKSKYPLLLDDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDGRSGYQSGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSWVLGSGSGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSWVLGSGSGVFKCG 650
Qy 601 IAVAPVSRWEYSDYVYTERYMGVLPEDNLDHYRNSVTVMRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYSDYVYTERYMGVLPEDNLDHYRNSVTVMRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 4
US-08-619-280A-3
; Sequence 3, Application US/08619280A
; Patent No. 5767242
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
;
US-08-619-280A-3
Query Match 100.0%; Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKXLSRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
Db 51 NTYRLKXLSRWISDHEYLKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
```

61	Qy	QOFILLEYNYKQWRHSYTASDYIDLNKRQLITEERIPNNTQWWTSPVGHKLAYVNN	120
		QOFILLEYNYKQWRHSYTASDYIDLNKRQLITEERIPNNTQWWTSPVGHKLAYVNN	170
121	Qy	DIYVKIEBPLPSYRLTWTKEDIIYNGITDWVYEEVFSAYSALMWSNPGTFLAYAQFND	180
		DIYVKIEBPLPSYRLTWTKEDIIYNGITDWVYEEVFSAYSALMWSNPGTFLAYAQFND	230
171	Db	DIYVKIEBPLPSYRLTWTKEDIIYNGITDWVYEEVFSAYSALMWSNPGTFLAYAQFND	240
181	Qy	TEVPLIEYSFYSDLESLOPKTVRPVYPKAGAVNPTKFFVNTDSLSVYTNATSQITAP	290
		TEVPLIEYSFYSDLESLOPKTVRPVYPKAGAVNPTKFFVNTDSLSVYTNATSQITAP	300
241	Qy	ASMLIGDHYLQDVWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST	350
		ASMLIGDHYLQDVWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST	410
301	Qy	TGWGRFRPSPBPHFTLDGNSFYKIIISNEEGYRHCYFOIDKKDCTFIHKGTWEVIGIEAL	460
		TGWGRFRPSPBPHFTLDGNSFYKIIISNEEGYRHCYFOIDKKDCTFIHKGTWEVIGIEAL	510
351	Db	TGWGRFRPSPBPHFTLDGNSFYKIIISNEEGYRHCYFOIDKKDCTFIHKGTWEVIGIEAL	560
361	Qy	TSDXLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFKEAKYYOL	610
		TSDXLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFKEAKYYOL	660
411	Db	TSDXLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFKEAKYYOL	710
421	Qy	RCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLPFIILNETKFWYQMTL	760
		RCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLPFIILNETKFWYQMTL	810
471	Db	RCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLPFIILNETKFWYQMTL	860
481	Qy	PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGSYQGDKI	910
		PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGSYQGDKI	960
531	Db	PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGSYQGDKI	1010
541	Qy	MHA1NRR1LGTTEVEDQIEAARQFQSKMGFVNDKRIAIWGSYGGYVTSMWLGSGSGVFKCG	1060
		MHA1NRR1LGTTEVEDQIEAARQFQSKMGFVNDKRIAIWGSYGGYVTSMWLGSGSGVFKCG	1110
591	Db	MHA1NRR1LGTTEVEDQIEAARQFQSKMGFVNDKRIAIWGSYGGYVTSMWLGSGSGVFKCG	1160
601	Qy	IAPVPSRWXYDYSVYTERYMG1PTPE1NLDHYRNSTVMSRAENPKQVEYLLIHGTADDN	1210
		IAPVPSRWXYDYSVYTERYMG1PTPE1NLDHYRNSTVMSRAENPKQVEYLLIHGTADDN	1260
651	Db	IAPVPSRWXYDYSVYTERYMG1PTPE1NLDHYRNSTVMSRAENPKQVEYLLIHGTADDN	1310
661	Qy	VHFQOQAQISKALVDVGVDFQAMWYTDDBDHGIASSTAHOHIIYTHMSHF1KQCFSLP	1360
		VHFQOQAQISKALVDVGVDFQAMWYTDDBDHGIASSTAHOHIIYTHMSHF1KQCFSLP	1410
711	Db	VHFQOQAQISKALVDVGVDFQAMWYTDDBDHGIASSTAHOHIIYTHMSHF1KQCFSLP	1460

RESULT 5

RESOLUTION 3
US-08-940-391-3

US-08-940-391-3
: Sequence 3. Application US/08940391

Sequence 3, APPLIC
Patent No. 5965373

FACEID NO: 3363373
GENERAL INFORMATION:

APPLICANT: Zimmermann, Rainer; park. John E.;

APPLICANT: Rettig, Wolfgang: Old, Lloyd J.

TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION

TITLE OF INVENTION: PROTEIN ALPHA, AND USES THEREOF

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: Felfe & Lynch

STREET: 805 Third A

; CITY: New York City

STATE: New York

; COUNTRY: USA

; ZIP: 10022

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage

```

;
COMPUTER:  IBM PS/2
OPERATING SYSTEM:  DOS

```

OPERATING SYSTEM

```

; SOFTWARE: wordperfect
: CURRENT APPLICATION DATA.

```

```

: : CURRENT APPLICATION DATA:
: : APPLICATION NUMBER: IIS

```

APPLICATION NUMBER: US/08/940,391
FILING DATE: 01-OCT-1997

REF ID: A66585

CLASSIFICATION: 330
PRIOR APPLICATION DATA:

UNIT 1

APPLICATION NUMBER: 08/619,280
 FILING DATE: 18-MARCH-1996
 APPLICATION NUMBER: 08/230,491
 FILING DATE: 20-APRIL-1994
 ATTORNEY/AGENT INFORMATION:
 NAME: Harison, No. 5965373man D.
 REGISTRATION NUMBER: 30,946
 REFERENCE/DOCKET NUMBER: LUD 5330.1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 688-9200
 TELEFAX: (212) 838-3884
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 766 amino acids
 TYPE: amino acid
 TOPOLOGY: linear

Query Match	100.0%;	Score 3877;	DB 1;	Length 766;	
Best Local Similarity	100.0%;	Pred. No. 0;			
Matches	716;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	NTYRLKLSRLWLSDEHYLYKQENNLLVFNAAEYGNSSVFLNENSTFDFBGHSINDYS	SPD	60	
Db	51	NTYRLKLSRLWLSDEHYLYKQENNLLVFNAAEYGNSSVFLNENSTFDFBGHSINDYS	SPD	110	
Qy	61	GQFILLYYNVKQWRHSYASDYIDLNRKQLITEERI PNNTQWVTSVPVGHKLAYVWNN	120		
Db	111	GQFILLYYNVKQWRHSYASDYIDLNRKQLITEERI PNNTQWVTSVPVGHKLAYVWNN	170		
Qy	121	DIYVKIEPNLPSYRIITWTGKEDI IYNGITDWTVEEVEFSAISALWSPNGTFLAYAQFND	180		
Db	171	DIYVKIEPNLPSYRIITWTGKEDI IYNGITDWTVEEVEFSAISALWSPNGTFLAYAQFND	230		
Qy	181	TEVPLLEYSPYSDESLOYPKTVRPVPKAGAVNPTVKFFVWNTDSLSVNTNATSIQITAP	240		
Db	231	TEVPLLEYSPYSDESLOYPKTVRPVPKAGAVNPTVKFFVWNTDSLSVNTNATSIQITAP	290		
Qy	241	ASMLIGDHYLCVWTATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVAROHIE MST	300		
Db	291	ASMLIGDHYLCVWTATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVAROHIE MST	350		
Qy	301	TGWGFRPSPHFPTLDGNSFYKII SNEEGYRHCYFQIDKDKCTFITKGTWIEVIGIEAL	360		
Db	351	TGWGFRPSPHFPTLDGNSFYKII SNEEGYRHCYFQIDKDKCTFITKGTWIEVIGIEAL	410		
Qy	361	TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYVSVSFSEAKYQYL	420		
Db	411	TSDYLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYVSVSFSEAKYQYL	470		
Qy	421	RCSPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKCLDFIILNETKFWYQML	480		
Db	471	RCSPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKCLDFIILNETKFWYQML	530		
Qy	481	PPHFDKSKYPLLLDVYAGPCKQADTVFRLNWTATYLASTENIIVASFDRGSGSYQGDKI	540		
Db	531	PPHFDKSKYPLLLDVYAGPCKQADTVFRLNWTATYLASTENIIVASFDRGSGSYQGDKI	590		
Qy	541	MHAINRLGTTFVEDQIEAARQFSKMGFVDNKRRIATWGSYGGYVTSWVLGSSGVPKCG	600		
Db	591	MHAINRLGTTFVEDQIEAARQFSKMGFVDNKRRIATWGSYGGYVTSWVLGSSGVPKCG	650		
Qy	601	IAPVPSRWEYDSVYTERYMGLPTEPNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	660		
Db	651	IAPVPSRWEYDSVYTERYMGLPTEPNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	710		
Qy	661	VHPQSQSAQTSKALVDYGVGVDFQAWYTTDEDHGIASSTAHQHIYTHMSHFIIKQCPSLP	716		
Db	711	VHPQSQSAQTSKALVDYGVGVDFQAWYTTDEDHGIASSTAHQHIYTHMSHFIIKQCPSLP	766		

RESIT.T 6

RESULTS
US-09-794-236-1

; Sequence 1, Application US/09794236
; Patent No. 6337069
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-Silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/794,236
; CURRENT FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-794-236-1

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTDFFGHISINDYSISPD 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
51 NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTDFFGHISINDYSISPD 110
Qy 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSAWSPNGTFLAYAQFND 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSAWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFVSDLSQYPKTVRPYPKAGAVNPVTKFFVNTDLSSTVNTATSIQTAP 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
231 TEVPLIEYSFVSDLSQYPKTVRPYPKAGAVNPVTKFFVNTDLSSTVNTATSIQTAP 290
Qy 241 ASMLIGHLYLDCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
291 ASMLIGHLYLDCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHTLDDGNSFYKIIISNEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
351 TGWVGRFRPSEPHTLDDGNSFYKIIISNEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 410
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCGPGLPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPCKLDFILNETKFWYQML 480
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
471 RCGPGLPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPCKLDFILNETKFWYQML 530
Qy 481 PPHFDKSKYPLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
531 PPHFDKSKYPLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQTEAARQFSKMGFVNDKRIAIWGSYGGYVTVSMVLGSGGVFKCG 600
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
591 MHAINRRLGTFEVEDQTEAARQFSKMGFVNDKRIAIWGSYGGYVTVSMVLGSGGVFKCG 650
Qy 601 IAVAPSRWEYSDSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKOVEYLLHGTADDN 660
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
651 IAVAPSRWEYSDSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKOVEYLLHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIKQCFSLP 716
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIKQCFSLP 766

RESULT 7
US-09-265-606-3

; Sequence 3, Application US/09265606
; Patent No. 6846910
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/265,606
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280
; FILING DATE: 18-MARCH-1996
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 6846910man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-3200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 766 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-09-265-606-3

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTDFFGHISINDYSISPD 60
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
51 NTYRLKLYSLRWISDHELYLKQENNILVFNAYGNSVFLNSTDFFGHISINDYSISPD 110
Qy 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 120
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSAWSPNGTFLAYAQFND 180
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSAWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFVSDLSQYPKTVRPYPKAGAVNPVTKFFVNTDLSSTVNTATSIQTAP 240
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
231 TEVPLIEYSFVSDLSQYPKTVRPYPKAGAVNPVTKFFVNTDLSSTVNTATSIQTAP 290
Qy 241 ASMLIGHLYLDCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 300
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
291 ASMLIGHLYLDCVDTWATQERISLOWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHTLDDGNSFYKIIISNEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 360
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||
351 TGWVGRFRPSEPHTLDDGNSFYKIIISNEGYRHCYFQIDKDKCTFTTKGTWEVIGIEAL 410
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db ||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 411 TSDYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDPIILINETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDPIILINETKFWQMIL 530
Qy 481 PPFFDKSKKYPILLDDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 540
Db 531 PPFFDKSKKYPILLDDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSWVLGSGSGGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSWVLGSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYDSVYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDSVYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 766

RESULT 8

US-09-949-016-10450
; Sequence 10450, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10450
; LENGTH: 775
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-10450

Query Match 100.0%; Score 3877; DB 2; Length 775;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NTYRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNENSTFDFEGHSINDYISPD 60
Db 60 NTYRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNENSTFDFEGHSINDYISPD 119
Qy 61 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 120
Db 120 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 179
Qy 121 DIYVKIEPNLPYSRIITWTGKEDIYNGITDWMYEEVFSAYSALWMSPNGTFLAYAQFND 180
Db 180 DIYVKIEPNLPYSRIITWTGKEDIYNGITDWMYEEVFSAYSALWMSPNGTFLAYAQFND 239
Qy 181 TEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 240
Db 240 TEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 299
Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 300 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 359
Qy 301 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIAL 360

Db 360 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKCTFITKGTWEVIGIAL 419
Qy 361 TSDYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 420 TSDYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 479
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDPIILINETKFWQMIL 480
Db 480 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDPIILINETKFWQMIL 539
Qy 481 PPFFDKSKKYPILLDDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 540
Db 540 PPFFDKSKKYPILLDDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 599
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSWVLGSGSGGVFKCG 600
Db 600 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKIAIWGWSYGGYVTSWVLGSGSGGVFKCG 659
Qy 601 IAVAPVSRWEYDSVYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 660
Db 660 IAVAPVSRWEYDSVYTERYMGILPTPEDNLDHYRSTVMSRAENFKQVEYLLIHGTADDN 719
Qy 661 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 716
Db 720 VHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 775

RESULT 9

US-10-002-593-6
; Sequence 6, Application US/10002593
; Patent No. 6586198
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTING
; ENZYME INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. 6586198 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NTYRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNENSTFDFEGHSINDYISPD 60
Db 51 NTYRLKLYSLRWISDHELYLKQENNLVFNABYGNSSVFLNENSTFDFEGHSINDYISPD 110
Qy 61 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 120
Db 111 GQFILLLEYNVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPYSRIITWTGKEDIYNGITDWMYEEVFSAYSALWMSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPYSRIITWTGKEDIYNGITDWMYEEVFSAYSALWMSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 240
Db 231 TEVPLIEYSFYSDESLOYPKTVRPVYPKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 290
Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350

QY 301 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVEIGIEAL 360
DB 351 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVEIGIEAL 410
QY 361 TSDLYLYISNEYKMGPGRNLYKIQLDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 420
DB 411 TSDLYLYISNEYKMGPGRNLYKIQLDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 470
QY 421 RCGSPGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 480
DB 471 RCGSPGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 530
QY 481 PPFDKSKYPLLLDVVYAGPCSKADTVFRLNWTYLASTENIIVASFDFGRSGYQGDKI 540
DB 531 PPFDKSKYPLLLDVVYAGPCSKADTVFRLNWTYLASTENIIVASFDFGRSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGGVFKCG 600
DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGGVFKCG 650
QY 601 IAVAPSRWEYDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 10
US-09-949-016-6146
; Sequence 6146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6146
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-6146

Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
DB 51 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
QY 61 GQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAAYVWNN 120
DB 111 GQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYLWSPNGTFLAYAQFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYLWSPNGTFLAYAQFND 230
QY 181 TEVPLIEYSFYSDSLQYPKTVRPVPKAGAVNPTVKFFVVTDSLSSVTNATSIQITAP 240

DB 231 TEVPLIEYSFYSDSLQYPKTVRPVPKAGAVNPTVKFFVVTDSLSSVTNATSIQITAP 290
QY 241 ASMLIGHYLCDDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRNCLVARQHIEMST 300
DB 291 ASMLIGHYLCDDVTWATQERISLOWLRRIQNYSVMDICDYDESSGRNCLVARQHIEMST 350
QY 301 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVEIGIEAL 360
DB 351 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVEIGIEAL 410
QY 361 TSDLYLYISNEYKMGPGRNLYKIQLDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 420
DB 411 TSDLYLYISNEYKMGPGRNLYKIQLDYTKVTCLSCELNPERCOYYSVSFSKEAKYYOL 470
QY 421 RCGSPGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 480
DB 471 RCGSPGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFFILNETKFWYQWIL 530
QY 481 PPFDKSKYPLLLDVVYAGPCSKADTVFRLNWTYLASTENIIVASFDFGRSGYQGDKI 540
DB 531 PPFDKSKYPLLLDVVYAGPCSKADTVFRLNWTYLASTENIIVASFDFGRSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGGVFKCG 600
DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGGVFKCG 650
QY 601 IAVAPSRWEYDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYDSVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 11
US-10-423-714-6
; Sequence 6, Application US/10423714
; Patent No. 6887679
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI
; FILE REFERENCE: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Atty Docket No. 6887679 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/423,714
; CURRENT FILING DATE: 2003-04-25
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-423-714-6

Query Match 99.8%; Score 3871; DB 2; Length 766;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
DB 51 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
QY 61 GQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAAYVWNN 120
DB 111 GQFILLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYLWSPNGTFLAYAQFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYLWSPNGTFLAYAQFND 230

181 TEVPLIEYSFYSDSLOQPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP 240
182 |
183 |
184 |
185 |
186 |
187 |
188 |
189 |
190 |
191 |
192 |
193 |
194 |
195 |
196 |
197 |
198 |
199 |
200 |
201 |
202 |
203 |
204 |
205 |
206 |
207 |
208 |
209 |
210 |
211 |
212 |
213 |
214 |
215 |
216 |
217 |
218 |
219 |
220 |
221 |
222 |
223 |
224 |
225 |
226 |
227 |
228 |
229 |
230 |
231 |
232 |
233 |
234 |
235 |
236 |
237 |
238 |
239 |
240 |
241 |
242 |
243 |
244 |
245 |
246 |
247 |
248 |
249 |
250 |
251 |
252 |
253 |
254 |
255 |
256 |
257 |
258 |
259 |
260 |
261 |
262 |
263 |
264 |
265 |
266 |
267 |
268 |
269 |
270 |
271 |
272 |
273 |
274 |
275 |
276 |
277 |
278 |
279 |
280 |
281 |
282 |
283 |
284 |
285 |
286 |
287 |
288 |
289 |
290 |
291 |
292 |
293 |
294 |
295 |
296 |
297 |
298 |
299 |
300 |
301 |
302 |
303 |
304 |
305 |
306 |
307 |
308 |
309 |
310 |
311 |
312 |
313 |
314 |
315 |
316 |
317 |
318 |
319 |
320 |
321 |
322 |
323 |
324 |
325 |
326 |
327 |
328 |
329 |
330 |
331 |
332 |
333 |
334 |
335 |
336 |
337 |
338 |
339 |
340 |
341 |
342 |
343 |
344 |
345 |
346 |
347 |
348 |
349 |
350 |
351 |
352 |
353 |
354 |
355 |
356 |
357 |
358 |
359 |
360 |
361 |
362 |
363 |
364 |
365 |
366 |
367 |
368 |
369 |
370 |
371 |
372 |
373 |
374 |
375 |
376 |
377 |
378 |
379 |
380 |
381 |
382 |
383 |
384 |
385 |
386 |
387 |
388 |
389 |
390 |
391 |
392 |
393 |
394 |
395 |
396 |
397 |
398 |
399 |
400 |
401 |
402 |
403 |
404 |
405 |
406 |
407 |
408 |
409 |
410 |
411 |
412 |
413 |
414 |
415 |
416 |
417 |
418 |
419 |
420 |
421 |
422 |
423 |
424 |
425 |
426 |
427 |
428 |
429 |
430 |
431 |
432 |
433 |
434 |
435 |
436 |
437 |
438 |
439 |
440 |
441 |
442 |
443 |
444 |
445 |
446 |
447 |
448 |
449 |
450 |
451 |
452 |
453 |
454 |
455 |
456 |
457 |
458 |
459 |
460 |
461 |
462 |
463 |
464 |
465 |
466 |
467 |
468 |
469 |
470 |
471 |
472 |
473 |
474 |
475 |
476 |
477 |
478 |
479 |
480 |
481 |
482 |
483 |
484 |
485 |
486 |
487 |
488 |
489 |
490 |
491 |
492 |
493 |
494 |
495 |
496 |
497 |
498 |
499 |
500 |
501 |
502 |
503 |
504 |
505 |
506 |
507 |
508 |
509 |
510 |
511 |
512 |
513 |
514 |
515 |
516 |
517 |
518 |
519 |
520 |
521 |
522 |
523 |
524 |
525 |
526 |
527 |
528 |
529 |
530 |
531 |
532 |
533 |
534 |
535 |
536 |
537 |
538 |
539 |
540 |
541 |
542 |
543 |
544 |
545 |
546 |
547 |
548 |
549 |
550 |
551 |
552 |
553 |
554 |
555 |
556 |
557 |
558 |
559 |
560 |
561 |
562 |
563 |
564 |
565 |
566 |
567 |
568 |
569 |
570 |
571 |
572 |
573 |
574 |
575 |
576 |
577 |
578 |
579 |
580 |
581 |
582 |
583 |
584 |
585 |
586 |
587 |
588 |
589 |
590 |
591 |
592 |
593 |
594 |
595 |
596 |
597 |
598 |
599 |
600 |
601 |
602 |
603 |
604 |
605 |
606 |
607 |
608 |
609 |
610 |
611 |
612 |
613 |
614 |
615 |
616 |
617 |
618 |
619 |
620 |
621 |
622 |
623 |
624 |
625 |
626 |
627 |
628 |
629 |
630 |
631 |
632 |
633 |
634 |
635 |
636 |
637 |
638 |
639 |
640 |
641 |
642 |
643 |
644 |
645 |
646 |
647 |
648 |
649 |
650 |
651 |
652 |
653 |
654 |
655 |
656 |
657 |
658 |
659 |
660 |
661 |
662 |
663 |
664 |
665 |
666 |
667 |
668 |
669 |
670 |
671 |
672 |
673 |
674 |
675 |
676 |
677 |
678 |
679 |
680 |
681 |
682 |
683 |
684 |
685 |
686 |
687 |
688 |
689 |
690 |
691 |
692 |
693 |
694 |
695 |
696 |
697 |
698 |
699 |
700 |
701 |
702 |
703 |
704 |
705 |
706 |
707 |
708 |
709 |
710 |
711 |
712 |
713 |
714 |
715 |
716 |
717 |
718 |
719 |
720 |
721 |
722 |
723 |
724 |
725 |
726 |
727 |
728 |
729 |
730 |
731 |
732 |
733 |
734 |
735 |
736 |
737 |
738 |
739 |
740 |
741 |
742 |
743 |
744 |
745 |
746 |
747 |
748 |
749 |
750 |
751 |
752 |
753 |
754 |
755 |
756 |
757 |
758 |
759 |
760 |
761 |
762 |
763 |
764 |
765 |
766 |
767 |
768 |
769 |
770 |
771 |
772 |
773 |
774 |
775 |
776 |
777 |
778 |
779 |
780 |
781 |
782 |
783 |
784 |
785 |
786 |
787 |
788 |
789 |
790 |
791 |
792 |
793 |
794 |
795 |
796 |
797 |
798 |
799 |
800 |
801 |
802 |
803 |
804 |
805 |
806 |
807 |
808 |
809 |
810 |
811 |
812 |
813 |
814 |
815 |
816 |
817 |
818 |
819 |
820 |
821 |
822 |
823 |
824 |
825 |
826 |
827 |
828 |
829 |
830 |
831 |
832 |
833 |
834 |
835 |
836 |
837 |
838 |
839 |
840 |
841 |
842 |
843 |
844 |
845 |
846 |
847 |
848 |
849 |
850 |
851 |
852 |
853 |
854 |
855 |
856 |
857 |
858 |
859 |
860 |
861 |
862 |
863 |
864 |
865 |
866 |
867 |
868 |
869 |
870 |
871 |
872 |
873 |
874 |
875 |
876 |
877 |
878 |
879 |
880 |
881 |
882 |
883 |
884 |
885 |
886 |
887 |
888 |
889 |
890 |
891 |
892 |
893 |
894 |
895 |
896 |
897 |
898 |
899 |
900 |
901 |
902 |
903 |
904 |
905 |
906 |
907 |
908 |
909 |
910 |
911 |
912 |
913 |
914 |
915 |
916 |
917 |
918 |
919 |
920 |
921 |
922 |
923 |
924 |
925 |
926 |
927 |
928 |
929 |
930 |
931 |
932 |
933 |
934 |
935 |
936 |
937 |
938 |
939 |
940 |
941 |
942 |
943 |
944 |
945 |
946 |
947 |
948 |
949 |
950 |
951 |
952 |
953 |
954 |
955 |
956 |
957 |
958 |
959 |
960 |
961 |
962 |
963 |
964 |
965 |
966 |
967 |
968 |
969 |
970 |
971 |
972 |
973 |
974 |
975 |
976 |
977 |
978 |
979 |
980 |
981 |
982 |
983 |
984 |
985 |
986 |
987 |
988 |
989 |
990 |
991 |
992 |
993 |
994 |
995 |
996 |
997 |
998 |
999 |
1000 |

RESULT 12

US-09-518-550-27
; Sequence 27, Application US/09518550
; Patent No. 6875851
; GENERAL INFORMATION:
; APPLICANT: TRAVIS, James
; APPLICANT: POTEMPA, Jan
; APPLICANT: BANBULA, Agnieszka
; TITLE OF INVENTION: PROLYL PEPTIDASES AND METHODS OF USE
; FILE REFERENCE: 235.00190101
; CURRENT APPLICATION NUMBER: US/09/518,550
; CURRENT FILING DATE: 2000-03-03
; PRIOR FILING DATE: 2000-03-03
; PRIOR FILING DATE: 1999-03-05
; PRIOR APPLICATION NUMBER: PCT/US00/05551
; PRIOR FILING DATE: 2000-03-03
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-518-550-27

Query Match 99.78; Score 3867; DB 2; Length 766;
Best Local Similarity 99.78; Pred. No. 0;
Matches 714; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 NTRRLKLYSRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDFEFGHSINDYSISPD 60
51 NTRRLKLYSRWISDHEYLKQENNILVFNAEYGNSSVFLNENSTFDFEFGHSINDYSISPD 110
61 GQFILLIYVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTSVPVGHKLAYVWNN 120
111 GQFILLIYVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWTSVPVGHKLAYVWNN 170

121 DIYVKLEPNLPSYRITWTGKEDI IYNGITDWTVEEVSAYSALWWSPNGTFLAYAQFND 180
171 DIYVKLEPNLPSYRITWTGKEDI IYNGITDWTVEEVSAYSALWWSPNGTFLAYAQFND 230
181 TEVPLIEYSFYSDSLOQPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP 240
231 TEVPLIEYSFYSDSLOQPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP 290
241 ASMLIGDHYLCVDTWATQERISLOWLRRIQNYSDMDICDYDESSGRWNCCLVARQHIEMST 300
291 ASMLIGDHYLCVDTWATQERISLOWLRRIQNYSDMDICDYDESSGRWNCCLVARQHIEMST 350
301 TGMVGRFRPSEPHFTLDGNSFYKII SNEEGYRHCICYFQIDKXDCCTFTTKGTWEVIGIEAL 360
351 TGMVGRFRPSEPHFTLDGNSFYKII SNEEGYRHCICYFQIDKXDCCTFTTKGTWEVIGIEAL 410
361 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
411 TSDLYLYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
421 RCGSGPLPLYTLHSSVNDKGLRVLEDNLSALDKMLQNVQMPSSKCLDPIILNETKFWYQML 480
471 RCGSGPLPLYTLHSSVNDKGLRVLEDNLSALDKMLQNVQMPSSKCLDPIILNETKFWYQML 530
481 PHFDDSKKYPILLDLYVAGCSQKADTVFRLNWTYLASTENIIVASFDGRSGSYQGDKI 540
531 PHFDDSKKYPILLDLYVAGCSQKADTVFRLNWTYLASTENIIVASFDGRSGSYQGDKI 590
541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVLSGSGGVFKCG 600
591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVLSGSGGVFKCG 650
601 IAVAPVSRWEYSDYVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
651 IAVAPVSRWEYSDYVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
661 VHFQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF1KQCFSLP 716
711 VHFQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHF1KQCFSLP 766

RESULT 13

PCT-US93-07923-11
; Sequence 11, Application PC/TUS9307923
; GENERAL INFORMATION:
; APPLICANT: Morimoto, Chikao
; APPLICANT: Schlossman, Stuart F.
; APPLICANT: Tanaka, Toshiaki
; TITLE OF INVENTION: HUMAN CD26 AND METHODS FOR USE
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: U.S.A.
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; COMPUTER: IBM PS/2 Model 50Z or 55SX
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
; SOFTWARE: WordPerfect (Version 5.0)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US93/07923
; FILING DATE: 19930819
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/934,162
; FILING DATE: 21-AUG-1992
; APPLICATION NUMBER: 07/832,211
; FILING DATE: 06-FEB-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Fraser, Janie K.

REGISTRATION NUMBER: 34,819
REFERENCE/DOCKET NUMBER: 00530/055002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 593
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
PCT-US93-07923-11

Query Match 76.0%; Score 2948; DB 4; Length 593;
Best Local Similarity 100.0%; Pred. No. 1.6e-252;
Matches 543; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTRYRLKLYSLRWISDHEVLYKQENNLVFNAEYGNSSVFLNSSTFDFGHSINDYSISPD 60
Db 51 NTRYRLKLYSLRWISDHEVLYKQENNLVFNAEYGNSSVFLNSSTFDFGHSINDYSISPD 110
Qy 61 GQFTLLEVNYKQWRHSYASYDYDLNKRQLITEERIPNNTQWTPSPVGHKLAYVNN 120
Db 111 GQFTLLEVNYKQWRHSYASYDYDLNKRQLITEERIPNNTQWTPSPVGHKLAYVNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSLWSPNGTFLAYAFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSLWSPNGTFLAYAFND 230
Qy 181 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGHYLCVDTWATQERISLOWLRRIQNYSDMDICDDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGHYLCVDTWATQERISLOWLRRIQNYSDMDICDDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWGRFRPSPHFTLQNSFYKISNEEGVYRHCYFQIDKCKCTFTKGTWVIGIEAL 360
Db 351 TGWGRFRPSPHFTLQNSFYKISNEEGVYRHCYFQIDKCKCTFTKGTWVIGIEAL 410
Qy 361 TSDYLYIYISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDYLYIYISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCGSGPLPLTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFILNETKFWYQML 480
Db 471 RCGSGPLPLTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFILNETKFWYQML 530
Qy 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 590
Qy 541 MHA 543
Db 591 MHA 593

RESULT 14
US-08-230-491A-2
Sequence 2, Application US/08230491A
Patent No. 5587299
GENERAL INFORMATION:
APPLICANT: Rettig, Wolfgang J.; Scanlan, Matthew J.;
APPLICANT: Garin-Chesa, Pilar; Old, Lloyd J.
TITLE OF INVENTION: ISOLATED NUCLEIC ACID MOLECULE CODING FOR
TITLE OF INVENTION: FIBROBLAST ACTIVATION PROTEIN ' AND USES
TITLE OF INVENTION: THEROP
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: FELFE & LYNCH
STREET: 805 THIRD AVENUE

CITY: NEW YORK
STATE: NEW YORK
COUNTRY: USA
ZIP: 10022
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE 3.5 inch 1.2 MB STORAGE
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: WORDPERFECT - ASC II
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/230,491A
FILING DATE: 20-APRIL-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Hanson, No. 5587299man D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LUD 330
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 760 amino acids
TYPE: amino acid
TOPOLOGY: linear
US-08-230-491A-2

Query Match 55.1%; Score 2138; DB 1; Length 760;
Best Local Similarity 53.3%; Pred. No. 1.7e-180;
Matches 383; Conservative 134; Mismatches 187; Indels 14; Gaps 7;
Qy 2 TYRLKLYSLRWISDHEVLYKQ-ENNILVFNAEYGNSSVFLNSSTFDFGHSIN--DYSIS 58
Db 51 TFSYKFFPNWISGQEVYHQSDANNVLYNIETGQSYTILSNRTM----KSNVNASNYGLS 106
Qy 59 PDGQFLLVYNYKQWRHSYASYDYDLNKRQLITEERIPNNTQWTPSPVGHKLAYVW 118
Db 107 PDQFVYLEDYKLRYSYATYIYDLSNGEFVRGNELPRPIQYLCWSPVGSKLAYVY 166
Qy 119 NNDIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSLWSPNGTFLAYAQF 178
Db 167 QNNIYKQRPDGPFPQITFNGRENKIPNGIPDMVYEEMLPTKYKALWSPNGFLAYAEF 226
Qy 179 NDTVEPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQIT 238
Db 227 NDKDIPVIAYSYGDE--QYPRINIPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQIT 281
Qy 239 APASMLIGHYLCVDTWATQERISLOWLRRIQNYSDMDICDDESSGRWNCCLVARQHIEM 298
Db 282 VPAMIASDDYFYSWLTWVTDVCLQWLKRVQNVSVLSICDPREDWQTDCKPTQEHIEE 341
Qy 299 STTGWGRFRPSPHFTLQNSFYKISNEEGVYRHCYFQIDKCKCTFTKGTWVIGIE 358
Db 342 SRTGWAGGFPVSRPVFSYDAISYKIFPSDKGDKYHIYKDTVENAQTITSGKEAINIF 401
Qy 359 ALTSDVLYIYISNEYKMGPGGRNLYKQLSDY--TKVTCLSCELNPERCQYYSVSFSKEAKY 417
Db 402 RVTDLSLFTSSNEFEYPGRRNIYRISIGSYPPSKKCVTCHLKRERQYITASFSDYAKY 461
Qy 418 YQLRCSGPGPLPLTLHSSVNDKGLRVLEDSALDKMLQNVQMPSSKKLDFILNETKFWYQ 477
Db 462 YALVCYCGPIPISTLHDGRDQEIKEENKELENALKNIQLPKKEIKLEVDITLWYK 521
Qy 478 MILPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGYQ 537
Db 522 MILPQFDRSKYPLLIQYVGGPCSVSRVFAVNMISYLASKEGMVIALVDGRGTAFQ 581
Qy 538 DKIMHAINRELGTFEVDDQIEARQPSKMGFVDNKRITAWGWSYGVVYTSMTVLGSGSVF 597
Db 582 DKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDKKRIAWGWSYGVVYTSMTVLGSGSVF 641
Qy 598 KCGIAPVAPSRWEYSDYVYTERVMGLPTPBDNLHDYRNSTVMGRNENFQVYELLIHGTA 657

```

Db      642 KCGIAPVSSWEYAYVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVVDYLLHNGTA 701
Qy      658 DDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIASTAHQIYTHMSHFHKQCFSL 715
Db      702 DDNVHFQOQAISKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTFLKQCFSL 758

RESULT 15
US-08-619-280A-2
; Sequence 2, Application US/08619280A
; Patent No. 5767242
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; APPLICANT: Rettig, Wolfgang; Old, Lloyd J.
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Felfe & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: PC-DOS
; SOFTWARE: Wordperfect
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/619,280A
; FILING DATE: 18-MARCH-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/230,491
; FILING DATE: 20-APRIL-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Hanson, No. 5767242man D.
; REGISTRATION NUMBER: 30,946
; REFERENCE/DOCKET NUMBER: LUD 5330.1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 688-9200
; TELEFAX: (212) 838-3884
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 760 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
US-08-619-280A-2

Query Match      55.1%; Score 2138; DB 1; Length 760;
Best Local Similarity 53.3%; Pred. No. 1.7e-180;
Matches 383; Conservative 134; Mismatches 187; Indels 14; Gaps 7;

Qy      2 TYRLKLYSLRWISDHEVLYKQ-ENNILVFNAEYGNSSVFLNSTFDFEGHSIN--DYSIS 58
Db      51 TFSYKTFPPNWSIQGEYLHQSADNNIVLYNIETGQSYTILSNRTM----KSNVNASNYGLS 106
Qy      59 PDGQFTILLEYNVYKQWRHSYASVDIYDLNKRQLITEERIPNNNTQWTVWSPVGHKLAYVW 118
Db      107 PDQFVYVLESYSLKWRYSYATYIYDLNNGEFGVGRNLPRIQVYLWSPVGSKLAYVY 166
Qy      119 NNDIYVKIENPLPSYRITWTGKEDIYNGITDWYBEEVFESAYSALWSPNGTFLAYAQF 178
Db      167 QNNIYLKQRPGDPFPQITFNGRENKIPNGIPDWYBEEMLPTKYALWWSPNGKFLAYAEF 226
Qy      179 NDFEVLIEYFYSDESLOYPKTVRVPYKAGAVNPTVFFVNTDSLSTVNTATSIQIT 238
Db      227 NDKDIPVIAYSYGDE--QYPRINIPYKAGAKNPVIRFIIDTTYPAYVGPQ---EVP 281
Qy      239 APASMLIGDHYLCVDTQWATERISLOWLRRIQYNSVMDICDYDESSGRWNCLVARQHIEM 298
Db      282 VPAMIASSDYFWSLWTVTDVRCVQLQWLKRVQVSVLSICDPREDWQTDWDCPKTQEHIE 341

```

```

Qy      299 STTGWGRFRPSEPHTLDGNSFYKLIISNEBGRHYCHYFQIDKKDCTFTTKGTWEVIGIE 358
Db      342 SRTGAGGFFVRPVSFYDAISYYKIFSDKGYGKHIHYIKDTVVENAIQITSGKWEAINIF 401
Qy      359 ALTSYLYYISNEYKMPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYVSFSFKEAKY 417
Db      402 RVTQDSLFSYSSNEFEYFGRRIYRISIGSYPPSKCVCTCHLKERCCQYVTASFSDYAKY 461
Qy      418 YQRCSGPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKFWYQ 477
Db      462 YALVCYGPPISTLHDGRDQEIKILEENKELENALKNIQLPKBEIKKLEVDIEITLWYK 521
Qy      478 MILPHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIVASFDRGSGYQG 537
Db      522 MILPPQFDRSKYPLLIQYVGPCSQSVRSVFAVNMISYLASKEGMVIALVDGRGTAFOG 581
Qy      538 DKIMHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRKRIALWGSYGGYVTSMLVSGSGGVF 597
Db      582 DKLLYAVYRKLGVYEVEDQITAVRKFIEMGFIDEKKRIALWGSYGGYVTSMLASGTGLF 641
Qy      598 KCGIAPVSRWEYSDSVYTERYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 657
Db      642 KCGIAPVSSWEYAYVYTERFMGLPTKDDNLEHYKNSTVMARAEYFRNVVDYLLHNGTA 701
Qy      658 DDNVHFQOQAISKALVDVGVDFQAMWYTDDEHGIASTAHQIYTHMSHFHKQCFSL 715
Db      702 DDNVHFQOQAISKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTFLKQCFSL 758

```

Search completed: January 27, 2006, 23:59:03
Job time : 30 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:57:48 ; Search time 67 Seconds
(without alignments)
4465.161 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTVRLKLYSLRWISDHEYLX.....AHQHYTHMSHFHKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA Main:
- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
 - 3: /cgn2_6/ptodata/1/pubpaa/US09_PUBCOMB.pep.*
 - 4: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
 - 6: /cgn2_6/ptodata/1/pubpaa/US11_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	3	US-09-265-606-3
2	3877	100.0	766	3	US-09-993-959-1
3	3877	100.0	766	4	US-10-295-027-590
4	3877	100.0	766	5	US-10-952-459-18
5	3877	100.0	766	5	US-10-631-467-565
6	3871	99.8	766	4	US-10-002-593-6
7	3871	99.8	766	4	US-10-165-603-7
8	3871	99.8	766	4	US-10-423-714-6
9	3871	99.8	766	4	US-10-295-027-922
10	3871	99.8	766	4	US-10-794-899-41
11	3871	99.8	766	6	US-11-041-674-6
12	3867	99.7	766	5	US-10-476-264-147
13	3867	99.7	766	5	US-10-770-712-1
14	3358.5	86.6	767	4	US-10-770-712-2
15	3347.5	86.3	767	4	US-10-165-603-4
16	3347.5	86.3	767	4	US-10-794-899-38
17	3340	86.1	760	5	US-10-631-467-1390
18	3324	85.7	760	5	US-10-770-712-3
19	2145	55.3	760	5	US-10-723-860-4171
20	2138	55.1	760	3	US-09-265-606-2
21	2138	55.1	760	4	US-10-177-293-136
22	2138	55.1	760	4	US-10-301-822-55
23	2138	55.1	760	5	US-10-884-070A-13
24	1276.5	32.9	504	4	US-10-072-012-863
25	1221	31.5	228	5	US-10-476-264-114
26	1217	31.4	746	5	US-10-476-264-69
27	1217	31.4	746	5	US-10-476-264-70

28	1217	31.4	746	5	US-10-476-264-72	Sequence 72, Appl
29	1217	31.4	789	5	US-10-476-264-71	Sequence 71, Appl
30	1217	31.4	796	3	US-09-976-674-5	Sequence 5, Appl
31	1217	31.4	796	3	US-09-870-133-2	Sequence 2, Appl
32	1217	31.4	796	4	US-10-160-501-5	Sequence 5, Appl
33	1217	31.4	796	5	US-10-982-512-5	Sequence 5, Appl
34	1217	31.4	796	5	US-10-433-757-2	Sequence 2, Appl
35	1217	31.4	796	5	US-10-476-264-68	Sequence 68, Appl
36	1217	31.4	796	5	US-10-476-264-106	Sequence 106, App
37	1217	31.4	796	5	US-10-476-264-110	Sequence 110, App
38	1217	31.4	796	5	US-10-476-264-142	Sequence 142, App
39	1217	31.4	796	5	US-10-476-264-146	Sequence 146, App
40	1217	31.4	797	5	US-10-476-264-36	Sequence 36, Appl
41	1211	31.2	798	4	US-10-210-130-120	Sequence 120, App
42	1198	30.9	743	4	US-10-363-937-4	Sequence 4, Appl
43	1198	30.9	743	6	US-11-187-040-4	Sequence 4, Appl
44	1196	30.8	706	3	US-09-976-674-41	Sequence 41, Appl
45	1196	30.8	706	5	US-10-982-512-41	Sequence 41, Appl

ALIGNMENTS

RESULT 1
US-09-265-606-3
; Sequence 3, Application US/09265606
; Patent No. US20020034789A1
; GENERAL INFORMATION:
; APPLICANT: Zimmermann, Rainer; Park, John E.;
; TITLE OF INVENTION: ISOLATED DIMERIC FIBROBLAST ACTIVATION PROTEIN
; TITLE OF INVENTION: ALPHA, AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Feife & Lynch
; STREET: 805 Third Avenue
; CITY: New York City
; STATE: New York
; COUNTRY: USA
; ZIP: 10022

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.5 inch, 2.0 MB storage
COMPUTER: IBM PS/2
OPERATING SYSTEM: PC-DOS
SOFTWARE: Wordperfect

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/265,606
FILING DATE:
CLASSIFICATION:

PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/619,280
FILING DATE: 18-MARCH-1996

APPLICATION NUMBER: 08/230,491
FILING DATE: 20-APRIL-1994
ATTORNEY/AGENT INFORMATION:

NAME: Hanson, No. US20020034789Alman D.
REGISTRATION NUMBER: 30,946
REFERENCE/DOCKET NUMBER: LJD 5330.1

TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 688-9200
TELEFAX: (212) 838-3884

INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 766 amino acids
TYPE: amino acid

TOPOLOGY: linear
US-09-265-606-3

Query Match 100.0%; Score 3877; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTVRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSFTDFGHSINDYSISPD 60

Db 51 NTRRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLNSTDFFEGHSINDYSISP 110
Qy 61 GQFILLEYNVVKQWRHSYTASDYIDLNKRQLITEERIIPNNTQVWTSVPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTASDYIDLNKRQLITEERIIPNNTQVWTSVPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRIITWTKGEDIINYGITDWTVEEVFSAYSALMWSNPGTFFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTKGEDIINYGITDWTVEEVFSAYSALMWSNPGTFFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSLSSTVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSLSSTVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLDCVTVWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLDCVTVWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYPQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYPQIDKKDCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 411 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMIL 480
Db 471 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMIL 530
Qy 481 PPHFDKSKYPPLLDVYAGCSOKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 540
Db 531 PPHFDKSKYPPLLDVYAGCSOKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSMLVSGSGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSMLVSGSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 766

RESULT 2
US-09-993-959-1
; Sequence 1, Application US/09993959
; Publication No. US20030165489A1
; GENERAL INFORMATION:
; APPLICANT: Grouzmann, Eric
; APPLICANT: Lacroix, Jean-silvain
; APPLICANT: Monod, Michel
; TITLE OF INVENTION: Method of Treating Rhinitis and Sinusitis
; FILE REFERENCE: 81985/276823
; CURRENT APPLICATION NUMBER: US/09/993,959
; CURRENT FILING DATE: 2001-11-27
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 1
; LENGTH: 766
; TYPE: PRN
; ORGANISM: Homo sapiens
US-09-993-959-1
Query Match 100.0%; Score 3877; DB 3; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NTRRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLNSTDFFEGHSINDYSISP 60

Db 51 NTRRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLNSTDFFEGHSINDYSISP 110
Qy 61 GQFILLEYNVVKQWRHSYTASDYIDLNKRQLITEERIIPNNTQVWTSVPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTASDYIDLNKRQLITEERIIPNNTQVWTSVPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRIITWTKGEDIINYGITDWTVEEVFSAYSALMWSNPGTFFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTKGEDIINYGITDWTVEEVFSAYSALMWSNPGTFFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSLSSTVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLESLOYPKTVRVPYPKAGAVNPTVKFFVNTDSLSSTVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLDCVTVWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLDCVTVWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYPQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKIIISNEEGYRHI CYPQIDKKDCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
Db 411 TSDLYYISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
Qy 421 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMIL 480
Db 471 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQMIL 530
Qy 481 PPHFDKSKYPPLLDVYAGCSOKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 540
Db 531 PPHFDKSKYPPLLDVYAGCSOKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSMLVSGSGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSMLVSGSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVYTERYMGFLPTPEDNLDRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 766

RESULT 3
US-10-295-027-590
; Sequence 590, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynnne, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and Methods for Modulators of Cancer
; FILE REFERENCE: 018501-012500US
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394

; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 590
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-590

Query Match 100.0%; Score 3877; DB 4; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTVRLKLYSLRWISDHEHYLYKQENNILVFNAYGNSSVFLENSTFDFGHSINDYSISPD 60
Db 51 NTVRLKLYSLRWISDHEHYLYKQENNILVFNAYGNSSVFLENSTFDFGHSINDYSISPD 110
Qy 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWVWTSVPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWVWTSVPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDI IYNGITDWTVEEVEVFSAYSA LWSVNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDI IYNGITDWTVEEVEVFSAYSA LWSVNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFVSDSLOYPKTVRVPYKAGAVNPTVKFVFNNTDLSLSSVTNATSIQTAP 240
Db 231 TEVPLIEYSFVSDSLOYPKTVRVPYKAGAVNPTVKFVFNNTDLSLSSVTNATSIQTAP 290
Qy 241 ASMLIGHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCIVARQHIEMST 300
Db 291 ASMLIGHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCIVARQHIEMST 350
Qy 301 TGVVGRFRPSEPFPFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGVVGRFRPSEPFPFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFILNETKFWYOMIL 480
Db 471 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFILNETKFWYOMIL 530
Qy 481 PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVFNKRIAIWGSYGGYVTVSMVLSGSGVVKPCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVFNKRIAIWGSYGGYVTVSMVLSGSGVVKPCG 650
Qy 601 IAVAPVSRWEYDVSVYTERYMGUPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVYTERYMGUPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIIKQCFSLP 716

; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 590
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-295-027-590

Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTVRLKLYSLRWISDHEHYLYKQENNILVFNAYGNSSVFLENSTFDFGHSINDYSISPD 60
Db 51 NTVRLKLYSLRWISDHEHYLYKQENNILVFNAYGNSSVFLENSTFDFGHSINDYSISPD 110
Qy 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWVWTSVPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWVWTSVPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDI IYNGITDWTVEEVEVFSAYSA LWSVNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDI IYNGITDWTVEEVEVFSAYSA LWSVNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFVSDSLOYPKTVRVPYKAGAVNPTVKFVFNNTDLSLSSVTNATSIQTAP 240
Db 231 TEVPLIEYSFVSDSLOYPKTVRVPYKAGAVNPTVKFVFNNTDLSLSSVTNATSIQTAP 290
Qy 241 ASMLIGHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCIVARQHIEMST 300
Db 291 ASMLIGHYLCVDTWATQERISLQWLRRIONYSVMDICDYDESSGRWNCIVARQHIEMST 350
Qy 301 TGVVGRFRPSEPFPFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 360
Db 351 TGVVGRFRPSEPFPFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWEVIGIEAL 410
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFILNETKFWYOMIL 480
Db 471 RCGSPGLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFILNETKFWYOMIL 530
Qy 481 PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVFNKRIAIWGSYGGYVTVSMVLSGSGVVKPCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVFNKRIAIWGSYGGYVTVSMVLSGSGVVKPCG 650
Qy 601 IAVAPVSRWEYDVSVYTERYMGUPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVYTERYMGUPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIIKQCFSLP 716

Db 711 VHFQSAQISKALVDGVDFQAMWYTDHGIASSTAHOHIYTHMSHFHKQCFSLP 766
|||||
RESULT 5
US-10-631-467-565
; Sequence 565, Application US/10631467
; Publication No. US20050208496A1
; GENERAL INFORMATION:
; APPLICANT: Genox Research Inc.
; TITLE OF INVENTION: Method for testing for bronchial asthma, or chronic obstructive p
; TITLE OF INVENTION: disease
; FILE REFERENCE: 3462.1005-000
; CURRENT APPLICATION NUMBER: US/10/631,467
; PRIOR FILING DATE: 2003-07-31
; PRIOR APPLICATION NUMBER: JP 2003-077212
; PRIOR FILING DATE: 2003-03-20
; PRIOR APPLICATION NUMBER: JP 2002-229312
; PRIOR FILING DATE: 2002-08-06
; NUMBER OF SEQ ID NOS: 2086
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 565
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-631-467-565

Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 5.8e-312;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTRYRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLENSTDFEFGHSINDYISPD 60
|||||
Db 51 NTRYRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLENSTDFEFGHSINDYISPD 110
|||||
Qy 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTSPVGHKLAYVWNN 120
|||||
Db 111 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTSPVGHKLAYVWNN 170
|||||
Qy 121 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 180
|||||
Db 171 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 230
|||||
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
|||||
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
|||||
Qy 241 ASMLIGDHVLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
|||||
Db 291 ASMLIGDHVLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
|||||
Qy 301 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHCYFQIDKKDCTFITKGTWEVIGIEAL 360
|||||
Db 351 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHCYFQIDKKDCTFITKGTWEVIGIEAL 410
|||||
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYVSFSKEAKYYQL 420
|||||
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYVSFSKEAKYYQL 470
|||||
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWYQML 480
|||||
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWYQML 530
|||||
Qy 481 PPHFDKSKYPLLLDVYAGPCQKADTVFRLNATYLASTENII VASFDGRGSGYQGDKI 540
|||||
Db 531 PPHFDKSKYPLLLDVYAGPCQKADTVFRLNATYLASTENII VASFDGRGSGYQGDKI 590
|||||
Qy 541 MHAINRRLGTFTVEVDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
|||||
Db 591 MHAINRRLGTFTVEVDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 650
|||||
Qy 601 IAVAPVSRWEYDVSVYTERYMGLPPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
|||||

Db 651 IAVAPVSRWEYDVSVYTERYMGLPPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDGVDFQAMWYTDHGIASSTAHOHIYTHMSHFHKQCFSLP 716
|||||
Db 711 VHFQSAQISKALVDGVDFQAMWYTDHGIASSTAHOHIYTHMSHFHKQCFSLP 766
|||||
RESULT 6
US-10-002-593-6
; Sequence 6, Application US/10002593
; Publication No. US20020137120A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. US20020137120A1 1242/48/2
; CURRENT APPLICATION NUMBER: US/10/002,593
; CURRENT FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-002-593-6

Query Match 99.8%; Score 3871; DB 4; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.8e-311;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTRYRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLENSTDFEFGHSINDYISPD 60
|||||
Db 51 NTRYRLKLSLRWISDHELYLKQENNLVFNABYGNSSVFLENSTDFEFGHSINDYISPD 110
|||||
Qy 61 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTSPVGHKLAYVWNN 120
|||||
Db 111 GQFILLEYNVVKWRHSYTSYDIYDLNKRQLITEERI PNNTQWTVTSPVGHKLAYVWNN 170
|||||
Qy 121 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 180
|||||
Db 171 DIYVKIEPNLPYRITWTGKEDI IYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 230
|||||
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 240
|||||
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSSLSSVTNATSIQITAP 290
|||||
Qy 241 ASMLIGDHVLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
|||||
Db 291 ASMLIGDHVLCVDTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
|||||
Qy 301 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHCYFQIDKKDCTFITKGTWEVIGIEAL 360
|||||
Db 351 TGWGRFRPSEPHTLDGNSFYKII SNEEGYRHCYFQIDKKDCTFITKGTWEVIGIEAL 410
|||||
Qy 361 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYVSFSKEAKYYQL 420
|||||
Db 411 TSDLYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYVSFSKEAKYYQL 470
|||||
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWYQML 480
|||||
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWYQML 530
|||||
Qy 481 PPHFDKSKYPLLLDVYAGPCQKADTVFRLNATYLASTENII VASFDGRGSGYQGDKI 540
|||||
Db 531 PPHFDKSKYPLLLDVYAGPCQKADTVFRLNATYLASTENII VASFDGRGSGYQGDKI 590
|||||
Qy 541 MHAINRRLGTFTVEVDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
|||||
Db 591 MHAINRRLGTFTVEVDQIEAARQFSKMGFVDNKRKRIAIWGSYGGYVTSMLVSGSGGVFKCG 650
|||||

Qy 601 IAVAPVSRWEYDVSYYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSYYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHQOQAQISKALVDVGVDFQAMWYTDDEHGIASHTAQHIIYTHMSHPFKQCFSLP 716
Db 711 VHQOQAQISKALVDVGVDFQAMWYTDDEHGIASHTAQHIIYTHMSHPFKQCFSLP 766

RESULT 7

US-10-165-603-7

; Sequence 7, Application US/10165603

; Publication No. US20030021792A1

; GENERAL INFORMATION:

; APPLICANT: Roben, Paul W.

; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE

; TITLE OF INVENTION: PROTEINS

; FILE REFERENCE: TPTECH.001A

; CURRENT APPLICATION NUMBER: US/10/165,603

; CURRENT FILING DATE: 2002-06-07

; PRIOR APPLICATION NUMBER: 60/297,021

; PRIOR FILING DATE: 2001-06-08

; PRIOR APPLICATION NUMBER: 60/305,117

; PRIOR FILING DATE: 2001-07-12

; NUMBER OF SEQ ID NOS: 33

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 7

; TYPE: PRT

; LENGTH: 766

; ORGANISM: Homo sapiens

US-10-165-603-7

Query Match 99.8%; Score 3871; DB 4; Length 766;

Best Local Similarity 99.9%; Pred. No. 1.8e-311;

Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTFDFGHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTFDFGHSINDYSISPD 110

Qy 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 120
Db 111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 170

Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVFSAYSAWNSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVFSAYSAWNSPNGTFLAYAQFND 230

Qy 181 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 290

Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIMST 300
Db 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIMST 350

Qy 301 TGVVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 360
Db 351 TGVVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 410

Qy 361 TSDYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470

Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFFILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFFILNETKFWQMIL 530

Qy 481 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590

Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKR IAIWGSYGGVYVTSWVLGSSGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKR IAIWGSYGGVYVTSWVLGSSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSYYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSYYTERYMGLEPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHQOQAQISKALVDVGVDFQAMWYTDDEHGIASHTAQHIIYTHMSHPFKQCFSLP 716
Db 711 VHQOQAQISKALVDVGVDFQAMWYTDDEHGIASHTAQHIIYTHMSHPFKQCFSLP 766

RESULT 8

US-10-423-714-6

; Sequence 6, Application US/10423714

; Publication No. US20030180828A1

; GENERAL INFORMATION:

; APPLICANT: Vanderbilt University

; APPLICANT: Brown, Nancy J.

; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTI

; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA

; FILE REFERENCE: Acty Docket No. US20030180828A1 1242/48/2

; CURRENT APPLICATION NUMBER: US/10/423,714

; CURRENT FILING DATE: 2003-04-25

; PRIOR APPLICATION NUMBER: 60/244,524

; PRIOR FILING DATE: 2000-10-31

; NUMBER OF SEQ ID NOS: 10

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 6

; TYPE: PRT

; LENGTH: 766

; ORGANISM: Homo sapiens

US-10-423-714-6

Query Match 99.8%; Score 3871; DB 4; Length 766;

Best Local Similarity 99.9%; Pred. No. 1.8e-311;

Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTFDFGHSINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHEHYLYKQENNLVFNAYGNSVFLNSTFDFGHSINDYSISPD 110

Qy 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 120
Db 111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERI PNTQVWTVSPVGHKLAYWNN 170

Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVFSAYSAWNSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWWYEEVFSAYSAWNSPNGTFLAYAQFND 230

Qy 181 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 290

Qy 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIMST 300
Db 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIMST 350

Qy 301 TGVVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 360
Db 351 TGVVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHCYFQIDKDCCTFTTKGTWEVIGIEAL 410

Qy 361 TSDYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 411 TSDYLYIISNEYKMGPGGRNLKYIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYYQL 470

Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFFILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFFILNETKFWQMIL 530

Qy 481 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 540
Db 531 PPHFDKSKYPLLLDDVYAGCSQKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI 590

Db 531 PPHFDKSKKYPDLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
Qy 541 MHAINRLGTFFVEDQIEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 600
Db 591 MHAINRLGTFFVEDQIEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 650
Qy 601 IAVAPSRWEYVDSVYTERYMGSLPTPDNLDHYRNSVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPSRWEYVDSVYTERYMGSLPTPDNLDHYRNSVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDHGHASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDHGHASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 9

US-10-295-027-922
; Sequence 922, Application US/10295027
; Publication No. US20030232350A1
; GENERAL INFORMATION:
; APPLICANT: Afar, Daniel
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsberg, Wendy M.
; APPLICANT: Gish, Kurt C.
; APPLICANT: Glynn, Richard
; APPLICANT: Hevezi, Peter A.
; APPLICANT: Mack, David H.
; APPLICANT: Murray, Richard
; APPLICANT: Watson, Susan R.
; APPLICANT: Eos Biotechnology, Inc.
; TITLE OF INVENTION: Methods of Diagnosis of Cancer, Compositions and
; FILE REFERENCE: 018501-0125000S
; CURRENT APPLICATION NUMBER: US/10/295,027
; CURRENT FILING DATE: 2002-11-13
; PRIOR APPLICATION NUMBER: US 09/663,733
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: US 60/350,666
; PRIOR FILING DATE: 2001-11-13
; PRIOR APPLICATION NUMBER: US 60/335,394
; PRIOR FILING DATE: 2001-11-15
; PRIOR APPLICATION NUMBER: US 60/332,464
; PRIOR FILING DATE: 2001-11-21
; PRIOR APPLICATION NUMBER: US 60/334,393
; PRIOR FILING DATE: 2001-11-29
; PRIOR APPLICATION NUMBER: US 60/340,376
; PRIOR FILING DATE: 2001-12-14
; PRIOR APPLICATION NUMBER: US 60/347,211
; PRIOR FILING DATE: 2002-01-08
; PRIOR APPLICATION NUMBER: US 60/347,349
; PRIOR FILING DATE: 2002-01-10
; PRIOR APPLICATION NUMBER: US 60/355,250
; PRIOR FILING DATE: 2002-02-08
; PRIOR APPLICATION NUMBER: US 60/356,714
; PRIOR FILING DATE: 2002-02-13
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1386
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 922
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-295-027-922

Query Match 99.8%; Score 3871; DB 4; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.8e-311;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NTRRLKLYSRWISDHEYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYISPD 60
Db 51 NTRRLKLYSRWISDHEYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWS PVGHKLA YVWNN 120

Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWS PVGHKLA YVWNN 170
Qy 121 DIYVKIEPMLPSYRITWTCKEDIIYNGITDWYVEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPMLPSYRITWTCKEDIIYNGITDWYVEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYDESLOYPKTVRVPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYDESLOYPKTVRVPYPKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHLYCDVTWATQERISLOWLRRIONYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHLYCDVTWATQERISLOWLRRIONYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360
Db 351 TGVVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 410
Qy 361 TSDYLYYISNEYKMGPPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYQOL 420
Db 411 TSDYLYYISNEYKMGPPGGRNLYKIQIDYTKVTCLSCELNPERCOYYSVSFSKEAKYQOL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWTQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPKSLDFIILNETKFWTQMIL 530
Qy 481 PPHFDKSKKYPDLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 540
Db 531 PPHFDKSKKYPDLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDGRSGYQGDKI 590
Qy 541 MHAINRLGTFFVEDQIEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 600
Db 591 MHAINRLGTFFVEDQIEAARQSKMGFVNDKRIAIWGSYGGYVTSWVLGSGGVFKCG 650
Qy 601 IAVAPSRWEYVDSVYTERYMGSLPTPDNLDHYRNSVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPSRWEYVDSVYTERYMGSLPTPDNLDHYRNSVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDHGHASSTAHOIYTHMSHFIKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDHGHASSTAHOIYTHMSHFIKQCFSLP 766

RESULT 10

US-10-794-899-41
; Sequence 41, Application US/10794899
; Publication No. US20040146516A1
; GENERAL INFORMATION:
; APPLICANT: Utah Ventures
; TITLE OF INVENTION: Lumen-Exposed Molecules and Methods Targeted Delivery
; FILE REFERENCE: 27110-715
; CURRENT APPLICATION NUMBER: US/10/794,899
; CURRENT FILING DATE: 2004-03-05
; NUMBER OF SEQ ID NOS: 121
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 41
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-794-899-41

Query Match 99.8%; Score 3871; DB 4; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.8e-311;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 NTRRLKLYSRWISDHEYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYISPD 60
Db 51 NTRRLKLYSRWISDHEYLKQENNILVFNAYGNSVFLNSTFDFGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWS PVGHKLA YVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWS PVGHKLA YVWNN 170

QY 121 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 240
DB 231 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 300
DB 291 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 350
QY 301 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 360
DB 351 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 410
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 420
DB 411 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 470
QY 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVMPKSKLDFIILNETKFWYQ 480
DB 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVMPKSKLDFIILNETKFWYQ 530
QY 481 PPFDKSKYPLLLDDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRGSGYQGDKI 540
DB 531 PPFDKSKYPLLLDDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRGSGYQGDKI 590
QY 541 MHAINRLGTPEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSWVLGSGGVKFCG 600
DB 591 MHAINRLGTPEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSWVLGSGGVKFCG 650
QY 601 IAVAPSRWEYSDYVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYSDYVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 11

US-11-041-674-6
; Sequence 6, Application US/11041674
; Publication No. US20050181468A1
; GENERAL INFORMATION:
; APPLICANT: Vanderbilt University
; APPLICANT: Brown, Nancy J.
; TITLE OF INVENTION: BIOLOGICAL MARKERS AND DIAGNOSTIC TESTS FOR ANGIOTENSIN CONVERTIN
; TITLE OF INVENTION: INHIBITOR AND VASOPEPTIDASE INHIBITOR ASSOCIATED ANGIOEDEMA
; FILE REFERENCE: Acty Docket No. 1242/48/2/2/2
; CURRENT APPLICATION NUMBER: US/11/041,674
; CURRENT FILING DATE: 2005-01-24
; PRIOR APPLICATION NUMBER: 60/244,524
; PRIOR FILING DATE: 2000-10-31
; PRIOR APPLICATION NUMBER: 10/002,593
; PRIOR FILING DATE: 2001-10-31
; PRIOR APPLICATION NUMBER: 10/423,714
; PRIOR FILING DATE: 2003-4-25
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 6
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-11-041-674-6

Query Match 99.8%; Score 3871; DB 6; Length 766;
Best Local Similarity 99.9%; Pred. No. 1.8e-311;
Matches 715; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSSTFDFGHSINDYSISPD 60

DB 51 NTYRLKLYSLRWISDHELYLKQENNILVFNAEYGNSSVFLNSSTFDFGHSINDYSISPD 110
QY 61 GQFTLELYNVVQWRHSYATSYDIYDLNKKQLITEERI PNTTQWVTHSPVGHKLAYVWNN 120
DB 111 GQFTLELYNVVQWRHSYATSYDIYDLNKKQLITEERI PNTTQWVTHSPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 240
DB 231 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNTPKFFVNTDLSSTVNTNATSIQITAP 290
QY 241 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 300
DB 291 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 350
QY 301 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 360
DB 351 TGWGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFQIDKKDCTFITKGTWVIGIEAL 410
QY 361 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 420
DB 411 TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQ 470
QY 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVMPKSKLDFIILNETKFWYQ 480
DB 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDSALDKMLQNVMPKSKLDFIILNETKFWYQ 530
QY 481 PPFDKSKYPLLLDDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRGSGYQGDKI 540
DB 531 PPFDKSKYPLLLDDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRGSGYQGDKI 590
QY 541 MHAINRLGTPEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSWVLGSGGVKFCG 600
DB 591 MHAINRLGTPEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGVYVTSWVLGSGGVKFCG 650
QY 601 IAVAPSRWEYSDYVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYSDYVYTERYNGLTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 12

US-10-476-264-147
; Sequence 147, Application US/10476264
; Publication No. US20050123910A1
; GENERAL INFORMATION:
; APPLICANT: Cookson, William Osmond Charles Michael
; APPLICANT: Moffat, Miriam Fleur
; APPLICANT: Allen, Maxine
; APPLICANT: Lench, Nick
; TITLE OF INVENTION: Enzyme and SNP marker for disease
; FILE REFERENCE: 16721-002US1
; CURRENT APPLICATION NUMBER: US/10/476,264
; CURRENT FILING DATE: 2003-10-24
; PRIOR APPLICATION NUMBER: PCT/GB02/01887
; PRIOR FILING DATE: 2002-04-24
; PRIOR APPLICATION NUMBER: GB0110044.5
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0110046.0
; PRIOR FILING DATE: 2001-04-24
; PRIOR APPLICATION NUMBER: GB0124594.3
; PRIOR FILING DATE: 2001-10-12
; PRIOR APPLICATION NUMBER: GB0124575.2
; PRIOR FILING DATE: 2001-10-12
; NUMBER OF SEQ ID NOS: 421
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 147
; LENGTH: 766
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-476-264-147

Query Match      99.7%; Score 3867; DB 5; Length 766;
Best Local Similarity 99.7%; Pred. No. 3.9e-311;
Matches 714; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTRYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSSTPDEFGHSINDYISPD 60
Db 51 NTRYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSSTPDEFGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHYLDCVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLDCVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFITKTGWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFITKTGWEVIGIEAL 410
Qy 361 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKEYQL 420
Db 411 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKEYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 530
Qy 481 PPFPDSSKKYPLLLDVTYAGCSOKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 540
Db 531 PPFPDSSKKYPLLLDVTYAGCSOKADIVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYYSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYYSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSSAQISKALVDVGVDFQAMWYTTDEDHGIASSTAHOHIYTHMSHFHKQCFSLP 716
Db 711 VHFQSSAQISKALVDVGVDFQAMWYTTDEDHGIASSTAHOHIYTHMSHFHKQCFSLP 766
```

```
RESULT 13
US-10-770-712-1
; Sequence 1, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSC12.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 766
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-770-712-1

Query Match      99.7%; Score 3867; DB 5; Length 766;
Best Local Similarity 99.7%; Pred. No. 3.9e-311;
Matches 714; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NTRYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSSTPDEFGHSINDYISPD 60
Db 51 NTRYRLKLYSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSSTPDEFGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERI PNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRIITWTGKEDIYNGITDWTVEEVEFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHYLDCVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLDCVTWATQERISLOWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFITKTGWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFITKTGWEVIGIEAL 410
Qy 361 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKEYQL 420
Db 411 TSDLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKEYQL 470
Qy 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 480
Db 471 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDFIILNETKFWQMIL 530
Qy 481 PPFPDSSKKYPLLLDVTYAGCSOKADTVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 540
Db 531 PPFPDSSKKYPLLLDVTYAGCSOKADIVFRLNWTATYLASTENIIIVASFDFGRSGGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVSGSGGVFKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRRIAIWGSYGGYVTSMLVSGSGGVFKCG 650
Qy 601 IAVAPVSRWEYYSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYYSVYTERYMGILPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSSAQISKALVDVGVDFQAMWYTTDEDHGIASSTAHOHIYTHMSHFHKQCFSLP 716
Db 711 VHFQSSAQISKALVDVGVDFQAMWYTTDEDHGIASSTAHOHIYTHMSHFHKQCFSLP 766
```

```
RESULT 14
US-10-770-712-2
; Sequence 2, Application US/10770712
; Publication No. US20050170333A1
; GENERAL INFORMATION:
; APPLICANT: Vojdani, Aristo
; TITLE OF INVENTION: IDENTIFICATION OF ETIOLOGY OF AUTISM
; FILE REFERENCE: IMSC12.008A
; CURRENT APPLICATION NUMBER: US/10/770,712
; CURRENT FILING DATE: 2004-02-03
; NUMBER OF SEQ ID NOS: 133
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus
```

US-10-770-712-2

Query Match 86.6%; Score 3358.5; DB 5; Length 767;
Best Local Similarity 84.8%; Pred. No. 5.4e-269;
Matches 609; Conservative 51; Mismatches 55; Indels 3; Gaps 1;

```
QY 1 NTVRLKLSLRLWSDHEYLKQENNLVFNAYGNSVFLNENSTFDFGHSINDYSISPD 60
DB 49 NTFRVKYSLSRWSDSEYLKQENNLVFNAYGNSVFLNENSTFDFGHSINDYSISPD 108
QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVGHKLAYVWNN 120
DB 109 RLFVLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVGHKLAYVWNN 168
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 180
DB 169 DIYVKIEPNLPSYRITWTGKEDIYNGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 228
QY 181 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVWNTDSLSVNTNATSIQTAP 240
DB 229 TGVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVWNTDSLSVNTNATSIQTAP 288
QY 241 ASMLIGHYLCVMTWATQERISLQWLRRIQNYVMDICDYDESSGRWNCVLAHQHIEWST 300
DB 289 ASVTTGDHYLCVMTWATQERISLQWLRRIQNYVMDICDYDESSGRWNCVLAHQHIEWST 348
QY 301 TGVWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKD---CTFITKGTWEVIGI 357
DB 349 TGVWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKD---CTFITKGTWEVIGI 408
QY 358 EALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVCLSCELNPERCOYYSVSFSKEAKY 417
DB 409 EALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVCLSCELNPERCOYYSVSFSKEAKY 468
QY 418 YQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQVMPKSKLDFFILNETRFWYQ 477
DB 469 YQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQVMPKSKLDFFILNETRFWYQ 528
QY 478 MILPPHFDKSKYPLLLDVYAGPCQKADTVFRLNWTATYLASTENIIVASFDFGRGSGYQ 537
DB 529 MILPPHFDKSKYPLLLDVYAGPCQKADTVFRLNWTATYLASTENIIVASFDFGRGSGYQ 588
QY 538 DKIMHAINRRLGTFEVDEQIEAARQFSKMGFVNDKRIAIGWSYGGYVTSWVLGSGGVF 597
DB 589 DKIMHAINRRLGTFEVDEQIEAARQFSKMGFVNDKRIAIGWSYGGYVTSWVLGSGGVF 648
QY 598 KCGIAPVPSRWWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 657
DB 649 KCGIAPVPSRWWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 708
QY 658 DDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSL 715
DB 709 DDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSL 766
```

RESULT 15

US-10-165-603-4
; Sequence 4, Application US/10165603
; Publication No. US20030021792A1
; GENERAL INFORMATION:
; APPLICANT: Roben, Paul W.
; TITLE OF INVENTION: TISSUE-SPECIFIC ENDOTHELIAL MEMBRANE
; TITLE OF INVENTION: PROTEINS
; FILE REFERENCE: TPECH.001A
; CURRENT APPLICATION NUMBER: US/10/165,603
; PRIOR FILING DATE: 2002-06-07
; PRIOR APPLICATION NUMBER: 60/297,021
; PRIOR FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: 60/305,117
; PRIOR FILING DATE: 2001-07-12
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 4
; LENGTH: 767
; TYPE: PRT
; ORGANISM: Rattus norvegicus
US-10-165-603-4

Query Match 86.3%; Score 3347.5; DB 4; Length 767;
Best Local Similarity 84.5%; Pred. No. 4.4e-268;
Matches 607; Conservative 52; Mismatches 56; Indels 3; Gaps 1;

```
QY 1 NTVRLKLSLRLWSDHEYLKQENNLVFNAYGNSVFLNENSTFDFGHSINDYSISPD 60
DB 49 NTFRVKYSLSRWSDSEYLKQENNLVFNAYGNSVFLNENSTFDFGHSINDYSISPD 108
QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVGHKLAYVWNN 120
DB 109 RLFVLLEYVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTSVGHKLAYVWNN 168
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 180
DB 169 DIYVKIEPNLPSYRITWTGKEDIYNGITDMWYEEVFSAYSALWSPNGTFLAYAQFND 228
QY 181 TEVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVWNTDSLSVNTNATSIQTAP 240
DB 229 TGVPLIEYSFYSDESLOYPKTVRPYPKAGAVNPTVKFFVWNTDSLSVNTNATSIQTAP 288
QY 241 ASMLIGHYLCVMTWATQERISLQWLRRIQNYVMDICDYDESSGRWNCVLAHQHIEWST 300
DB 289 ASVTTGDHYLCVMTWATQERISLQWLRRIQNYVMDICDYDESSGRWNCVLAHQHIEWST 348
QY 301 TGVWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKD---CTFITKGTWEVIGI 357
DB 349 TGVWGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKD---CTFITKGTWEVIGI 408
QY 358 EALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVCLSCELNPERCOYYSVSFSKEAKY 417
DB 409 EALTSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVCLSCELNPERCOYYSVSFSKEAKY 468
QY 418 YQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQVMPKSKLDFFILNETRFWYQ 477
DB 469 YQLRCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQVMPKSKLDFFILNETRFWYQ 528
QY 478 MILPPHFDKSKYPLLLDVYAGPCQKADTVFRLNWTATYLASTENIIVASFDFGRGSGYQ 537
DB 529 MILPPHFDKSKYPLLLDVYAGPCQKADTVFRLNWTATYLASTENIIVASFDFGRGSGYQ 588
QY 538 DKIMHAINRRLGTFEVDEQIEAARQFSKMGFVNDKRIAIGWSYGGYVTSWVLGSGGVF 597
DB 589 DKIMHAINRRLGTFEVDEQIEAARQFSKMGFVNDKRIAIGWSYGGYVTSWVLGSGGVF 648
QY 598 KCGIAPVPSRWWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 657
DB 649 KCGIAPVPSRWWEYDVSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 708
QY 658 DDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSL 715
DB 709 DDNVHFQOQAISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSL 766
```

Search completed: January 28, 2006, 00:04:55
Job time : 73 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:07:15 ; Search time 135 Seconds
(without alignments)
2330.335 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877
Sequence: 1 NTVRLKLYSLRWISDHEYLX.....AHQHIVTHMGHFKQCFSLP 716

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_21.*
1: Geneseqp1980s.*
2: Geneseqp1990s.*
3: Geneseqp2000s.*
4: Geneseqp2001s.*
5: Geneseqp2002s.*
6: Geneseqp2003as.*
7: Geneseqp2003bs.*
8: Geneseqp2004s.*
9: Geneseqp2005s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3877	100.0	759	2 AAR54612	Aar54612 Delta3-9
2	3877	100.0	766	2 AAR40909	Aar40909 Sequence
3	3877	100.0	766	5 ABB08991	Abb08991 Human dip
4	3877	100.0	766	5 AAG78417	Aag78417 Human dip
5	3877	100.0	766	7 ADD27855	Add27855 Human dip
6	3877	100.0	766	7 ADD46934	Add46934 Human pro
7	3877	100.0	766	7 ADN39272	Adn39272 Cancer/an
8	3877	100.0	766	8 ADJ83981	Adj83981 Human ful
9	3877	100.0	766	8 ADJ75313	Adj75313 Marker ge
10	3877	100.0	766	8 ADO19398	Ado19398 Human pro
11	3877	100.0	766	8 ADO19806	Ado19806 Human pro
12	3877	100.0	766	8 ADO71612	Ado71612 Amino aci
13	3877	100.0	766	8 ADO71644	Ado71644 Amino aci
14	3877	100.0	766	8 ABM80355	Abm80355 Tumour-as
15	3877	100.0	766	8 ADP54458	Adp54458 Human pro
16	3877	100.0	766	8 ADU06688	Adu06688 Novel bro
17	3877	100.0	766	8 ADV25525	Adv25525 Human dip
18	3877	100.0	766	9 ADY15161	Ady15161 PRO polyp
19	3877	100.0	766	9 ADY16580	Ady16580 PRO polyp
20	3877	100.0	766	9 ADZ14038	Adz14038 Human dip
21	3877	100.0	766	9 AEB94223	Aeb94223 CD26/dipe
22	3871	99.8	736	8 ADO40240	Ado40240 Human dpp
23	3871	99.8	766	5 ABG61910	Abg61910 Prostate
24	3871	99.8	766	5 AAO15555	Aao15555 Human dip

25	3871	99.8	766	6 ABP56700	Abp56700 Human liv
26	3871	99.8	766	7 ADD14045	Add14045 Human src
27	3871	99.8	766	7 ADN39604	Adn39604 Cancer/an
28	3871	99.8	766	8 ADO19400	Ado19400 Human pro
29	3867	99.7	766	6 ABP55629	Abp55629 Human dpp
30	3867	99.7	766	8 ADQ80365	Adq80365 Dipeptidy
31	3867	99.7	766	9 AEB77579	Aeb77579 Human dip
32	3866	99.7	766	2 AAR54611	Aar54611 Native CD
33	3779	97.5	739	2 AAR54613	Aar54613 Delta24-3
34	3441	88.8	688	8 ADO71642	Ado71642 Amino aci
35	3358.5	86.6	767	9 AEB77580	Aeb77580 Rat dipep
36	3355.5	86.5	767	3 AAB11748	Aab11748 Rat dipep
37	3354.5	86.5	767	7 ADD46932	Add46932 Rat Prote
38	3347.5	86.3	767	6 ABP56699	Abp56699 Rat liver
39	3340	86.1	760	8 ADJ76138	Adj76138 Marker ge
40	3340	86.1	760	8 ADO71646	Ado71646 Amino aci
41	3340	86.1	760	9 AEB94226	Aeb94226 Mouse CD2
42	3324	85.7	760	9 AEB77581	Aeb77581 Mouse dip
43	2948	76.0	593	2 AAR40916	Aar40916 Sequence
44	2948	76.0	593	2 AAR54614	Aar54614 Delta594-
45	2145	55.3	723	9 AEB94227	Aeb94227 Human sol

ALIGNMENTS

RESULT 1
AAR54612
ID AAR54612 standard; protein; 759 AA.
XX AC AAR54612;
XX AC
DT 25-MAR-2003 (revised)
DT 09-DEC-1994 (first entry)
XX
DE Delta3-9 CD26.
XX
KW Human; T cell activation antigen; CD26; analogues; deletion; soluble;
KW signal peptidase; immune-stimulating; response-stimulating; AIDS;
KW immunosuppression; AIDS-related complex.
XX Homo sapiens.
XX
FH Key Location/Qualifiers
FT Misc-difference 2..3
FT /note= "Position of delta3-9 deletion"
XX
PN WO9409132-A1.
XX
PD 28-APR-1994.
XX
PF 19-AUG-1993; 93WO-US007923.
XX
PR 21-AUG-1992; 92US-00934162.
XX (DAND) DANA FARBER CANCER INST INC.
XX
PI Morimoto C, Schlossman S, Tanaka T;
XX
PF 1994-151317/18.
XX
PT Polypeptide fragments and analogues of CD26 and encoding nucleic acid -
PT usefull for stimulating immune response, e.g. for treatment of AIDS to
PT counteract immunosuppressive drug, and as vaccine adjuvant.
XX
PS Claim 3; Page 49-52; 85pp; English.
XX
CC The sequences given in AAR54612-14 represents analogues of the human T
CC cell activation antigen CD26 which have internal deletions. The analogues
CC pref. lack residues 3-9 or 24-34. These analogues are soluble under
CC physiological conditions and lack enough amino acid residues to render
CC them susceptible to cleavage by signal peptidase. The peptide fragments
CC and analogues are useful as immune or response- stimulating therapeutics,

CC eg. they may be used for treatment of disease conditions characterised by
CC immunosuppression, eg. AIDS or AIDS-related complex, other virally or
CC environmentally-induced conditions, and certain congenital immune
CC deficiencies. The peptides can be employed to increase immune function
CC which has been impaired by use of immunosuppressive drugs, such as certain
CC chemotherapeutic drugs. (Updated on 25-MAR-2003 to correct PN field.)
XX
SQ Sequence 759 AA;

Query Match 100.0%; Score 3877; DB 2; Length 759;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTDFEFGHSINDYSISPD 60
Db 44 NTYRLKLSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTDFEFGHSINDYSISPD 103

QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVWSPVGHKLAYVWNN 120
Db 104 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVWSPVGHKLAYVWNN 163

QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 164 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 223

QY 181 TEVPLIEYSYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP 240
Db 224 TEVPLIEYSYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP 283

QY 241 ASMLIGHYLCYDVTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 300
Db 284 ASMLIGHYLCYDVTWATQERISLQWLRRIQNYSDICDYDESSGRWNCCLVARQHIEMST 343

QY 301 TGWGRFRPSEPFTLDGNSFYKIIISNEGYRHICYFQIDKDCCTFITKGTWEVIGIEAL 360
Db 344 TGWGRFRPSEPFTLDGNSFYKIIISNEGYRHICYFQIDKDCCTFITKGTWEVIGIEAL 403

QY 361 TSDYLYIISNEYKMGPGGRNLYKIQISDYTKVTCLSCELNPERCQYYSVFSKEAKYQL 420
Db 404 TSDYLYIISNEYKMGPGGRNLYKIQISDYTKVTCLSCELNPERCQYYSVFSKEAKYQL 463

QY 421 RCGSGGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKLDFTILNETKFWQMIL 480
Db 464 RCGSGGLPLYTLHSSVNDKGLRVLEDSALDKMLQNVQMPSKLDFTILNETKFWQMIL 523

QY 481 PPFDKSKKYPDLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASFDCGRSGYQGDKI 540
Db 524 PPFDKSKKYPDLLDVYAGCSQKADTVFRLNWTATYLASTENIIVASFDCGRSGYQGDKI 583

QY 541 MHAINRLGTFFVEDQIEAARQFSKMGFVNDKRIATGWISYGGYVTSWVLGSGGVFKCG 600
Db 584 MHAINRLGTFFVEDQIEAARQFSKMGFVNDKRIATGWISYGGYVTSWVLGSGGVFKCG 643

QY 601 IAVAPVSRWEYSDVSYTERVMGLPTPEDNLHDYRNSTVMSRAENKQVEVLLIHTGTADDN 660
Db 644 IAVAPVSRWEYSDVSYTERVMGLPTPEDNLHDYRNSTVMSRAENKQVEVLLIHTGTADDN 703

QY 661 VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIKQCFSLP 716
Db 704 VHFQQAQISKALVDGVDFQAMWYTDDEHGIASSTAQHIIYTHMSHFIKQCFSLP 759

RESULT 2
AAR40909
ID AAR40909 standard; protein; 766 AA.
XX
AC AAR40909;
XX
DT 25-MAR-2003 (revised)
DT 05-FEB-1994 (first entry)
XX
DE Sequence encoded by human CD26 cDNA.
XX

KW Human T cell activation antigen; monoclonal antibody Tal.
XX
OS Homo sapiens.
XX
PH Key Location/Qualifiers
FT 7..28 /label= hydrophobic
FT 29..323 /label= N-terminal glycosylated region of extracellular domain
FT /note= "8 sites for N-linked glycans"
FT 324..551 /label= Cysteine rich region of extracellular domain
FT /note= "1 N-linked glycosylation site"
FT 552..766 /label= C-terminal region of extracellular domain
FT /note= "1 N-linked glycosylation site & 1 catalytic site"
FT 627..631 /label= active site of serine protease/esterase
FT /note= "fits the consensus sequence GXSGX"
XX

PN WO9316102-A1.
XX
XX 19-AUG-1993.
XX
XX 09-APR-1992; 92WO-US002892.
XX
XX 06-FEB-1992; 92US-00832211.
XX
XX (DAND) DANA FARBER CANCER INST INC.
XX
XX Morimoto C, Schlossman SF, Tanaka T;
XX
XX WPI; 1993-272827/34.
XX
XX N-PSDB; AAQ46089.
XX
XX Polypeptide fragments of CD26 - are capable of disrupting binding of CD45
XX and CD26 and thus interfering with T-cell activation.
XX
XX Disclosure; Page 39-43; 73pp; English.

XX C26 is a human T cell activation antigen originally identified by its
XX reactivity with the MAb Tal. C26 cDNA library was constructed from human
XX PHA-activated T cells using the CDM7vector. The hydrophobic N-terminal of
XX the predicted CD26 polypeptide has the characteristics of a signal
XX sequence of the type II membrane protein, which is reinforced by the
XX observation that potential N-glycosylation sites are located in the
XX carboxy side of the hydrophobic core. Therefore the N-terminal 6 AAs are
XX predicted to be cytoplasmic, the next 22 AAs are predicted to transverse
XX the cytoplasmic membrane, and the 738 C-terminal AAs constitute the
XX predicted extracellular domain. (Updated on 25-MAR-2003 to correct PN
XX field.)

SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTDFEFGHSINDYSISPD 60
Db 51 NTYRLKLSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTDFEFGHSINDYSISPD 110

QY 61 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVWSPVGHKLAYVWNN 170

QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWSPNGTFLAYAQFND 230

QY 181 TEVPLIEYSYDESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQTAP 240
XX

Db 231 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNPTKFFVVTNDSLSSVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 350
Qy 301 TGVVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICFYQIDKKDCTFITKGTWVEIGIEAL 360
Db 351 TGVVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICFYQIDKKDCTFITKGTWVEIGIEAL 410
Qy 361 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWYQML 480
Db 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWYQML 530
Qy 481 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 540
Db 531 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 650
Qy 601 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFOAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFOAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 3

ABB08991
ID ABB08991 standard; protein; 766 AA.
AC ABB08991;
XX
XX 19-JUN-2002 (first entry)
XX
DE Human dipeptidyl peptidase IV.
XX
KW Human; dipeptidyl peptidase IV; antiasthmatic; antiallergic;
KW antinflammatory.
XX
OS Homo sapiens.
XX
XX US6337069-B1.
XX
XX 08-JAN-2002.
XX
XX 28-FEB-2001; 2001US-00794236.
XX
XX 28-FEB-2001; 2001US-00794236.
XX
XX (BMRA-) BMRA CORP BV.
XX
XX Grouzmann E, Lacroix J, Monod M;
XX WPI; 2002-163235/21.
XX

XX Treating a patient for mucosal inflammation associated with rhinitis,
XX sinusitis or both, by intranasally administering a peptidase that cleaves
XX at Xaa-Pro sequences, to the patient.
XX
XX Disclosure; Col 9-14; 13pp; English.
XX
XX Thus invention relates to the treating of a patient for mucosal
XX inflammation associated with rhinitis or sinusitis, comprising
XX intranasally administering a peptidase. The peptidase is considered

CC antiasthmatic, antiallergic and antinflammatory in its action. The
CC peptidase cleaves at Xaa-Pro sequences and is useful for treating a
CC patient for mucosal inflammation associated with rhinitis or sinusitis,
CC which is the result of allergies or asthma. This sequence represents
CC human dipeptidyl peptidase IV
XX

SQ Sequence 766 AA;
Query Match 100.0%; Score 3877; DB 5; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTVRLKLYLRWISDHELYLKQENNLVFNAYGNSVFLNSFTDFGHSINDYSIPD 60
Db 51 NTVRLKLYLRWISDHELYLKQENNLVFNAYGNSVFLNSFTDFGHSINDYSIPD 110
Qy 61 GQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTVSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKWRHSYTYASYDIYDLNKRQLITEERIPNNTQVWTVSPVGHKLAYVWNN 170
Qy 121 DIVVKIEPNLPSYRITWTGKEDIYNGITDMVVEEYVFSAYLSALWSPNGTFLAYAQFND 180
Db 171 DIVVKIEPNLPSYRITWTGKEDIYNGITDMVVEEYVFSAYLSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNPTKFFVVTNDSLSSVTNATSIQTAP 240
Db 231 TEVPLIEYSFSDLSQPKTVRPYPKAGAVNPTKFFVVTNDSLSSVTNATSIQTAP 290
Qy 241 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATERISLOWLRRIQNSVMDICDYDESSGRWNCVLARQHIEMST 350
Qy 301 TGVVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICFYQIDKKDCTFITKGTWVEIGIEAL 360
Db 351 TGVVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICFYQIDKKDCTFITKGTWVEIGIEAL 410
Qy 361 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYIYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWYQML 480
Db 471 RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWYQML 530
Qy 481 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 540
Db 531 PPFDKSKKYPDLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 590
Qy 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 600
Db 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSWVLGSGSGVPKCG 650
Qy 601 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDYSVTYRYMGLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFOAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFOAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 4

AAG78417
ID AAG78417 standard; protein; 766 AA.
XX
XX AAG78417;
XX
XX 12-APR-2002 (first entry)
XX
XX Human dipeptidyl peptidase IV amino acid sequence.
XX
XX 21953 prolyl oligopeptidase; antibody; proline; endopeptidase; cancer;
KW cardiovascular disease; autoimmune disease; atopic allergy;
KW

Db	111	GQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN	170
Qy	121	DIYVKIEPNLPSYRIITWTGKEDIYINGITDWMYEEVFSAYSALWSPNGTFLAYAQFND	180
Db	171	DIYVKIEPNLPSYRIITWTGKEDIYINGITDWMYEEVFSAYSALWSPNGTFLAYAQFND	230
Qy	181	TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP	240
Db	231	TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP	290
Qy	241	ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDMIDCYDESSGRWNCILVARQHIE	300
Db	291	ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDMIDCYDESSGRWNCILVARQHIE	350
Qy	301	TGWGRFRPSPHFTLDGNSFYKLIISNEEGYRHCYFQIDKDKDCTFITKGTWVIGIEAL	360
Db	351	TGWGRFRPSPHFTLDGNSFYKLIISNEEGYRHCYFQIDKDKDCTFITKGTWVIGIEAL	410
Qy	361	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVFSKEAKYQOL	420
Db	411	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVFSKEAKYQOL	470
Qy	421	RCSGPGLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWYQIL	480
Db	471	RCSGPGLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFWYQIL	530
Qy	481	PPHFDKSKYPLLLDVYAGCSQKADTVFRLNATYLASTENIIVASFQGRSGYQGDKI	540
Db	531	PPHFDKSKYPLLLDVYAGCSQKADTVFRLNATYLASTENIIVASFQGRSGYQGDKI	590
Qy	541	MHAINRLRGTFEVEDQIEAARQFSKMGFVGNKRIATWGSYGGVVTSMVLGSGGVFKCG	600
Db	591	MHAINRLRGTFEVEDQIEAARQFSKMGFVGNKRIATWGSYGGVVTSMVLGSGGVFKCG	650
Qy	601	IAPVSRWEYDYSVYTERYMGFLPTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN	660
Db	651	IAPVSRWEYDYSVYTERYMGFLPTPEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN	710
Qy	661	VHFQOQAISKALVDGVDFQAMWYTDHEDHGIASSTAHQHYTHMSHFIKQCFSLP	716
Db	711	VHFQOQAISKALVDGVDFQAMWYTDHEDHGIASSTAHQHYTHMSHFIKQCFSLP	766
RESULT 7			
ADN39272			
ID	ADN39272 standard; protein; 766 AA.		
XX			
AC	ADN39272;		
DT	17-JUN-2004 (first entry)		
XX			
DE	Cancer/angiogenesis/fibrosis-related polypeptide, SEQ ID NO:590.		
XX			
KW	Human; differential expression; cancer; angiogenic disorder;		
KW	fibrotic disorder; psoriasis; ischaemia; heart disease; atherosclerosis;		
KW	inflammatory disease; autoimmune disease;		
KW	retinal neovascularisation syndrome; scarring; uterine fibroid;		
KW	detection; diagnosis; prognosis; drug screening; drug targeting;		
KW	wound healing; contraception; cytostatic; cardiant; immunomodulatory;		
KW	vulnerable; gene therapy; vaccine.		
XX			
OS	Homo sapiens.		
XX			
PN	WO2003042661-A2.		
XX			
PD	22-MAY-2003.		
XX			
PF	13-NOV-2002; 2002WO-US036810.		
XX			
PR	13-NOV-2001; 2001US-0350666P.		
PR	21-NOV-2001; 2001US-0332464P.		
PR	29-NOV-2001; 2001US-0334393P.		
PR	03-DEC-2001; 2001US-0335394P.		

PR	14-DEC-2001; 2001US-0340376P.				
PR	08-JAN-2002; 2002US-0347211P.				
PR	10-JAN-2002; 2002US-0347349P.				
PR	08-FEB-2002; 2002US-0352520P.				
PR	13-FEB-2002; 2002US-0356714P.				
PR	20-FEB-2002; 2002US-0359077P.				
PR	29-MAR-2002; 2002US-0368809P.				
PR	04-APR-2002; 2002US-0370110P.				
PR	12-APR-2002; 2002US-0372246P.				
PR	05-JUN-2002; 2002US-0386614P.				
PR	16-JUL-2002; 2002US-0396839P.				
PR	22-JUL-2002; 2002US-0397775P.				
PR	22-JUL-2002; 2002US-0397845P.				
PR	09-SEP-2002; 2002US-0409450P.				
XX	(EOSB-) EOS BIOTECHNOLOGY INC.				
XX					
PI	Afar D, Aziz N, Ginsburg WM, Gish KC, Glynn R, Hevezi PA;				
PI	Mack DH, Murray R, Watson SR, Wilson KE, Ziolknik A;				
XX	WPI: 2003-468649/44.				
DR	N-PSDB; ADN39271.				
XX					
PT	Determining the presence or absence of a pathological cell in a patient,				
PT	useful for diagnosing, prognosing or treating cancer, comprises detecting				
PT	a nucleic acid in a biological sample.				
XX					
PS	Claim 12; SEQ ID NO 590; 1385pp; English.				
XX					
CC	The invention relates to nucleic acids and proteins (ADN38683-ADN40064)				
CC	whose expression is upregulated or downregulated in specific cancers or				
CC	other diseases such as angiogenic or fibrotic disorders, and to methods				
CC	of determining the presence or absence of a pathological cell in a				
CC	patient by detecting a nucleic acid at least 80% identical to those of				
CC	the invention or by detecting a polypeptide of the invention. The				
CC	invention also relates to expression vectors and host cells comprising a				
CC	nucleic acid of the invention; antibodies which specifically bind a				
CC	polypeptide of the invention; use of such antibodies for drug targeting;				
CC	and methods of screening for modulators of activity or expression of the				
CC	polypeptides and nucleic acids. The nucleic acids, polypeptides, and				
CC	antibodies and methods are useful for diagnosing, prognosing and treating				
CC	cancer and other conditions such as psoriasis, ischaemia, heart disease,				
CC	atherosclerosis, inflammatory diseases, autoimmune diseases, retinal				
CC	neovascularisation syndromes, scarring and uterine fibroids. They may				
CC	also be useful in wound healing and in contraception. The present				
CC	sequence represents a polypeptide of the invention.				
XX					
SQ	Sequence 766 AA;				
Query Match 100.0%; Score 3877; DB 7; Length 766;					
Best Local Similarity 100.0%; Pred. No. 0;					
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	1	NTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTDFPGHNSINDYSISPD	60		
Db	51	NTYRLKLYSLRWISDHEYLKQENNLVFNAYGNSVFLNSTDFPGHNSINDYSISPD	110		
Qy	61	GQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN	120		
Db	111	GQFILLVYVQWRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN	170		
Qy	121	DIYVKIEPNLPSYRIITWTGKEDIYINGITDWMYEEVFSAYSALWSPNGTFLAYAQFND	180		
Db	171	DIYVKIEPNLPSYRIITWTGKEDIYINGITDWMYEEVFSAYSALWSPNGTFLAYAQFND	230		
Qy	181	TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP	240		
Db	231	TEVPLIEYSFYSDLSQYPKTVRPYPKAGAVNPTVKFFVNTDSLSSTNATSIQITAP	290		
Qy	241	ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDMIDCYDESSGRWNCILVARQHIE	300		
Db	291	ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDMIDCYDESSGRWNCILVARQHIE	350		

QY 301 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360
 DB 351 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 410
 QY 361 TSDLYYISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420
 DB 411 TSDLYYISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470
 QY 421 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKLDLFILNETKFWYQML 480
 DB 471 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKLDLFILNETKFWYQML 530
 QY 481 PPHFDKSKKYPDLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 540
 DB 531 PPHFDKSKKYPDLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 590
 QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVFKCG 600
 DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVFKCG 650
 QY 601 IAVAPVSRWEYDYSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 651 IAVAPVSRWEYDYSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDGVDFQAMWYTDHGHGASSTAHOHIYTHMSHFIKQCFSLP 716
 DB 711 VHFQSAQISKALVDGVDFQAMWYTDHGHGASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 8

ADJ83981
 ID ADJ83981 standard; protein; 766 AA.

XX AC ADJ83981;

DT 06-MAY-2004 (first entry)

XX Human full-length colon dipeptidyl peptidase IV (DPP4V) protein.

XX crystal; proteni co-ordinate data; dipeptidyl peptidase IV; DPP4V;
 KW immunomodulatory; antidiabetic; antinflammatory; neuroprotective;
 KW antithyroid; antirheumatic; antiarthritic; anti-HIV; cytostatic;
 KW immune response; diabetes; inflammation; multiple sclerosis;
 KW Grave's disease; chronic rheumatoid arthritis; AIDS; cancer; human;
 KW colon; enzyme.

XX Homo sapiens.

XX WO2004011640-A1.

XX 05-FEB-2004.

XX 28-JUL-2003; 2003WO-JP0009523.

XX 29-JUL-2002; 2002US-0398761P.

XX (TANA) TANABE SEIVAKU CO.

XX Hiramatsu H, Kyono K, Shima H, Sugiyama S;

XX WPI; 2004-156830/15.

XX N-PSDB; ADJ83980.

PT New crystal of dipeptidyl peptidase IV capable of analyzing its three-
 PT dimensional structure, useful for designing, identifying, evaluating or
 PT searching an effector of the dipeptidyl peptidase IV.

XX Claim 3; SEQ ID NO 2; 332pp; English.

XX The invention relates to a novel crystal of a dipeptidyl peptidase IV
 CC (DPP4V) which is sufficient to ensure a resolution capable of analysing
 CC its three-dimensional structure up to the side chain level by X-ray
 CC crystallographic structural analysis. The crystal of the invention

CC demonstrates immunomodulatory, antidiabetic, antinflammatory,
 CC neuroprotective, antithyroid, antirheumatic, antiarthritic, anti-HIV and
 CC cytostatic activities and may be useful for providing a three-dimensional
 CC structural coordinate as the information for designing, identifying, IV.
 CC The effector may be useful as a modulatory agent of immune response and
 CC as a therapeutic or prophylactic agent for diabetes, inflammation,
 CC multiple sclerosis, Grave's disease, chronic rheumatoid arthritis, AIDS
 CC or cancer. The current sequence is that of the human full-length colon
 CC dipeptidyl peptidase IV (DPP4V) protein of the invention.

XX Sequence 766 AA;

QY Query Match 100.0%; Score 3877; DB 8; Length 766;
 DB Best Local Similarity 100.0%; Pred. No. 0;
 Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHELYLKQENNLIVFNAEYGNSSVPLENSTFDFGHSINDYSISPD 60

DB 51 NTYRLKLYSLRWISDHELYLKQENNLIVFNAEYGNSSVPLENSTFDFGHSINDYSISPD 110

QY 61 GQFILLEYNVVKQMRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVYNN 120

DB 111 GQFILLEYNVVKQMRHSYTSYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVYNN 170

QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYSLMWSPNGTFLAYAQFND 180

DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFVSAYSLMWSPNGTFLAYAQFND 230

QY 181 TEVPLIEYSFYSDLESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 240

DB 231 TEVPLIEYSFYSDLESLOYPKTVRVPYKAGAVNPTVKFFVNTDLSSTVNTATSIQITAP 290

QY 241 ASMLIGDHVLCVDTWATQERISLOWLRRIQNSVMDICDYDESSGRWNCCLVARQHIEMST 300

DB 291 ASMLIGDHVLCVDTWATQERISLOWLRRIQNSVMDICDYDESSGRWNCCLVARQHIEMST 350

QY 301 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 360

DB 351 TGVGRFRPSPHFTLDGNSFYKIIISNEEGYRHCYFQIDKKDCTFTTKGTWEVIGIEAL 410

QY 361 TSDLYYISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 420

DB 411 TSDLYYISNEYKMGPGGRNLYKIQSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL 470

QY 421 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKLDLFILNETKFWYQML 480

DB 471 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSEKLDLFILNETKFWYQML 530

QY 481 PPHFDKSKKYPDLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 540

DB 531 PPHFDKSKKYPDLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDGRSGYQGDKI 590

QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVFKCG 600

DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRKRIAIWGSYGGVYTSWVLGSGGVFKCG 650

QY 601 IAVAPVSRWEYDYSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660

DB 651 IAVAPVSRWEYDYSVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710

QY 661 VHFQSAQISKALVDGVDFQAMWYTDHGHGASSTAHOHIYTHMSHFIKQCFSLP 716

DB 711 VHFQSAQISKALVDGVDFQAMWYTDHGHGASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 9

ADJ75313
 ID ADJ75313 standard; protein; 766 AA.

XX AC ADJ75313;

XX DT 20-MAY-2004 (first entry)

XX	Marker gene related amino acid sequence SEQ ID NO:565.
DE	
XX	
KW	bronchial asthma; chronic obstructive pulmonary disease;
KW	respiratory epithelial cell; interleukin-13; respiratory; antiasthmatic;
XX	gene therapy; marker.
OS	Homo sapiens.
XX	
PN	EPI394274-A2.
XX	
PD	03-MAR-2004.
XX	
XX	04-AUG-2003; 2003EP-00254857.
XX	
PR	06-AUG-2002; 2002JP-00229312.
PR	20-MAR-2003; 2003JP-00077212.
XX	
PA	(GENO-) GENOX RES INC.
XX	
PI	Ohtani N, Sugita Y, Yamaya M, Kubo H, Nagai H, Izuwara K;
XX	
DR	WPI; 2004-193155/19.
XX	
PT	Testing for bronchial asthma or chronic obstructive pulmonary disease by
PT	comparing the expression level of a marker gene in a biological sample
PT	from a subject with the expression level of the gene in a sample from a
PT	healthy subject.
XX	
PS	Example 11; SEQ ID NO 565; 241pp; English.
XX	
CC	The present invention describes a method of testing for bronchial asthma
CC	or chronic obstructive pulmonary disease. The method comprises
CC	determining the expression level of a marker gene in a biological sample
CC	from a subject, comparing the expression level determined with the
CC	expression level of the marker gene in a biological sample from a healthy
CC	subject, and judging whether the subject has bronchial asthma or chronic
CC	obstructive pulmonary disease. The marker gene comprises: (a) a group of
CC	genes (S1) whose expression levels increase when respiratory epithelial
CC	cells are stimulated with interleukin-13; or (b) a group of genes (S2)
CC	whose expression levels decrease when respiratory epithelial cells are
CC	stimulated with interleukin-13. Also described: (1) a reagent (I) for
CC	testing for bronchial asthma or chronic obstructive pulmonary disease;
CC	(2) a kit for screening for a candidate compound for a therapeutic agent
CC	to treat bronchial asthma or chronic obstructive pulmonary disease; (3)
CC	an animal model for bronchial asthma or chronic obstructive pulmonary
CC	disease; (4) an inducer that induces bronchial asthma in a mouse; (5) a
CC	method for producing an animal model for bronchial asthma or chronic
CC	obstructive pulmonary disease; (6) a therapeutic agent for bronchial
CC	asthma or chronic obstructive pulmonary disease, comprising the compound,
CC	a marker gene or an antisense nucleic acid corresponding to a portion of
CC	the marker gene, a ribozyme, a polynucleotide that suppresses the
CC	expression of the gene through an RNAi effect or an antibody recognising
CC	a protein encoded by a marker gene; and (7) a DNA chip for testing for
CC	bronchial asthma or a chronic obstructive pulmonary disease, on which a
CC	probe has been immobilised to assay a marker gene. (I) has respiratory
CC	and antiasthmatic activities, and can be used in gene therapy. The method
CC	is useful for testing for or screening for a therapeutic agent for
CC	bronchial asthma or chronic obstructive pulmonary disease. The present
CC	sequence is used in the exemplification of the present invention.
XX	
SQ	Sequence 766 AA;
	Query Match 100.0%; Score 3877; DB 8; Length 766;
	Best Local Similarity 100.0%; Pred. No. 0;
	Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 NTYRLKLSRWISDHEVLYKQNNILVFNAYEYGNSSVFLNFSFDFEGHSINDYSIPD 60
DB	51 NTYRLKLSRWISDHEVLYKQNNILVFNAYEYGNSSVFLNFSFDFEGHSINDYSIPD 110
QY	61 GQFILLEVNVKQWRHSYASDYIDLNKRQLITEERIPNNNTQWTVSPVGHKLAYVNN 120

XX Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX N-PSDB; ADO19397.
DR WPI; 2004-420067/39.
DR N-PSDB; ADO19397.
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX Claim 7; SEQ ID NO 328; 1731pp; English.
PS
XX The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTVRLKLYSLRWISDHVLYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 60
DB |||||
QY 51 NTVRLKLYSLRWISDHVLYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 110
DB |||||
QY 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
DB |||||
QY 111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 170
DB |||||
QY 121 DIYKIEPNLPSYRITWTGKEDIYNGITDWYEEVFSAYSAALWSPNGTFLAYAQFND 180
DB |||||
QY 171 DIYKIEPNLPSYRITWTGKEDIYNGITDWYEEVFSAYSAALWSPNGTFLAYAQFND 230
DB |||||
QY 181 TEVPLIEYFYSDESLOYPKTVRVPYPKAGAVNPTVFFVNTDLSLSSVNTATSIQITAP 240
DB |||||
QY 231 TEVPLIEYFYSDESLOYPKTVRVPYPKAGAVNPTVFFVNTDLSLSSVNTATSIQITAP 290
DB |||||
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 300
DB |||||
QY 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 350
DB |||||
QY 301 TGWVGRFRPSPHPTLDCNSFYKILISNEBGRHICYFQIDKKOCTFTTKGTWEVIGIEAL 360
DB |||||
QY 351 TGWVGRFRPSPHPTLDCNSFYKILISNEBGRHICYFQIDKKOCTFTTKGTWEVIGIEAL 410
DB |||||
QY 361 TSDLYLYISNEYKMGPCGRNLKYQLSDYTKVTCCLSCENLPERCOYYSVSFSKEAKYYQL 420
DB |||||
QY 411 TSDLYLYISNEYKMGPCGRNLKYQLSDYTKVTCCLSCENLPERCOYYSVSFSKEAKYYQL 470
DB |||||
QY 421 RCGSGPLPLTLHSSVNDKGLRVLEDNSALDKMLQNVQMPCKLDFILNETKFWYQML 480
DB |||||
QY 471 RCGSGPLPLTLHSSVNDKGLRVLEDNSALDKMLQNVQMPCKLDFILNETKFWYQML 530
DB |||||
QY 481 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDCRGSGYQGDKI 540
DB |||||
QY 531 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDCRGSGYQGDKI 590
DB |||||
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWLVGSSGVFKCG 600
DB |||||
QY 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWLVGSSGVFKCG 650
DB |||||

QY 601 IAVAPVSRWEYDSVYTERYMGVLTPTEDNLHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB |||||
QY 651 IAVAPVSRWEYDSVYTERYMGVLTPTEDNLHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
DB |||||
QY 661 VHFQOSQAISKALVDVGVDFQAMWYTDDEHGIIASSSTAHOHIYTHMSHFQKCFSLP 716
DB |||||
QY 711 VHFQOSQAISKALVDVGVDFQAMWYTDDEHGIIASSSTAHOHIYTHMSHFQKCFSLP 766
DB |||||

RESULT 11
ADO19806
ID ADO19806 standard; protein; 766 AA.
XX
AC ADO19806;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human PRO polypeptide #365.
XX
KW Human; PRO; immune related disorder; systemic lupus erythematosus;
KW rheumatoid arthritis; osteoarthritis; juvenile chronic arthritis;
KW systemic sclerosis; Sjogren's syndrome; vasculitis; sarcoidosis;
KW autoimmune haemolytic anaemia; autoimmune thrombocytopenia; thyroiditis;
KW diabetes mellitus; renal disease; demyelinating disease;
KW central nervous system; peripheral nervous system;
KW demyelinating polyneuropathy; Guillain-Barre syndrome;
KW chronic inflammatory demyelinating polyneuropathy.
XX
OS Homo sapiens.
XX
PN WO2004043361-A2.
XX
PD 27-MAY-2004.
XX
PF 06-NOV-2003; 2003WO-US035268.
XX
PR 08-NOV-2002; 2002US-0425235P.
XX
PA (GETH) GENENTECH INC.
XX
PI Fong S, Dennis K, Clark H, Chiu H, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
XX WPI; 2004-420067/39.
DR N-PSDB; ADO19805.
XX
XX Novel PRO polypeptide e.g., PRO69614, PRO71106, or PRO86388 useful for
PT treating an immune related disorder such as systemic lupus erythematosus,
PT rheumatoid arthritis, osteoarthritis, juvenile chronic arthritis or
PT spondyloarthropathy.
XX
PS Claim 7; SEQ ID NO 730; 1731pp; English.
XX
CC The invention relates to human PRO polypeptides and the polynucleotides
CC encoding them. The polypeptides and polynucleotides are useful for
CC treating and diagnosing immune related disorders in mammals. The immune
CC related disorders include systemic lupus erythematosus, rheumatoid
CC arthritis, osteoarthritis, juvenile chronic arthritis, systemic
CC sclerosis, Sjogren's syndrome, vasculitis, sarcoidosis, autoimmune
CC haemolytic anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes
CC mellitus, immune-mediated renal disease, demyelinating diseases of the
CC central or peripheral nervous system, demyelinating polyneuropathy,
CC Guillain-Barre syndrome and chronic inflammatory demyelinating
CC polyneuropathy. This sequence represents a human PRO polypeptide of the
CC invention.
XX
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTVRLKLYSLRWISDHVLYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 60
DB |||||
QY 51 NTVRLKLYSLRWISDHVLYKQENNLVFNAYGSSVFLNSTDFFGHGSHINDYSISPD 110
DB |||||
QY 61 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
DB |||||
QY 111 GQFILLEYNVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 170
DB |||||
QY 121 DIYKIEPNLPSYRITWTGKEDIYNGITDWYEEVFSAYSAALWSPNGTFLAYAQFND 180
DB |||||
QY 171 DIYKIEPNLPSYRITWTGKEDIYNGITDWYEEVFSAYSAALWSPNGTFLAYAQFND 230
DB |||||
QY 181 TEVPLIEYFYSDESLOYPKTVRVPYPKAGAVNPTVFFVNTDLSLSSVNTATSIQITAP 240
DB |||||
QY 231 TEVPLIEYFYSDESLOYPKTVRVPYPKAGAVNPTVFFVNTDLSLSSVNTATSIQITAP 290
DB |||||
QY 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 300
DB |||||
QY 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNSYVMDICDYDESSGRWNCCLVARQHIEMST 350
DB |||||
QY 301 TGWVGRFRPSPHPTLDCNSFYKILISNEBGRHICYFQIDKKOCTFTTKGTWEVIGIEAL 360
DB |||||
QY 351 TGWVGRFRPSPHPTLDCNSFYKILISNEBGRHICYFQIDKKOCTFTTKGTWEVIGIEAL 410
DB |||||
QY 361 TSDLYLYISNEYKMGPCGRNLKYQLSDYTKVTCCLSCENLPERCOYYSVSFSKEAKYYQL 420
DB |||||
QY 411 TSDLYLYISNEYKMGPCGRNLKYQLSDYTKVTCCLSCENLPERCOYYSVSFSKEAKYYQL 470
DB |||||
QY 421 RCGSGPLPLTLHSSVNDKGLRVLEDNSALDKMLQNVQMPCKLDFILNETKFWYQML 480
DB |||||
QY 471 RCGSGPLPLTLHSSVNDKGLRVLEDNSALDKMLQNVQMPCKLDFILNETKFWYQML 530
DB |||||
QY 481 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDCRGSGYQGDKI 540
DB |||||
QY 531 PPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASFDCRGSGYQGDKI 590
DB |||||
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWLVGSSGVFKCG 600
DB |||||
QY 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSWLVGSSGVFKCG 650
DB |||||

Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTFDFEGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFTITKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFTITKGTWEVIGIEAL 410
Qy 361 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQOL 420
Db 411 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQOL 470
Qy 421 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVMPSKKLDFFILNETKFWYQML 480
Db 471 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVMPSKKLDFFILNETKFWYQML 530
Qy 481 PPHFDKSKKYPILLDVIYAGCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 540
Db 531 PPHFDKSKKYPILLDVIYAGCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 590
Qy 541 MHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRRIALWGSYGGYVTSWVLGSGGVFKCG 600
Db 591 MHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRRIALWGSYGGYVTSWVLGSGGVFKCG 650
Qy 601 IAVAPVSRWEYSDVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYSDVYTERYMGFLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 716
Db 711 VHFQOQAISKALVDVGVDFQAMWYTDDEHGIASSTAHQHIYTHMSHFIKQCFSLP 766

RESULT 12
AD071612
ID AD071612 standard; protein; 766 AA.

XX
AC AD071612;
XT 26-AUG-2004 (first entry)
DE Amino acid sequence of a human CD26 protein.
XX
KW CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
KW dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
KW cell cycle arrest; tumour; tumour necrosis; immune response; human.
XX
OS Homo sapiens.
XX
PN WO2004045497-A2.
XX
PD 03-JUN-2004.
XX
PF 15-MAY-2003; 2003WO-US015499.
XX
PR 17-MAY-2002; 2002US-0381606P.
XX
PA (TEXA) UNIV TEXAS SYSTEM.

XX
PI Dang NH, Morimoto C;
XX
DR WPI; 2004-420511/39.
XX N-PSDB; AD071611, AD071613.
XX
PT Use of a CD26 composition, and a chemotherapeutic and/or a
PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
PT cycle arrest, killing a cancer cell, treating cancer, or inducing tumor
PT regression or tumor necrosis.
XX
PS Claim 23; Page 151-153; 182pp; English.
XX
CC The specification describes a CD26 composition which, in conjunction with
CC chemotherapeutic or radiotherapeutic agents, is used for the treatment
CC and prevention of cancers. Expression of CD26 enhances the sensitivity of
CC the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
CC is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
CC topoisomerase II inhibitor. The CD26 composition of the invention is
CC useful for inhibiting the growth of a cell, inducing cell cycle arrest in
CC a cell, killing a cancer cell, potentiating the effect of a
CC chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
CC inducing or enhancing apoptosis of a cancer cell, treating cancer, or
CC inducing tumour regression or tumour necrosis. The CD26 composition is
CC further useful for increasing topoisomerase II expression in a cell, for
CC activating an antigen-presenting cell, or for potentiating immune
CC responses of an animal. The present sequence represents a CD26 protein,
CC and is encoded by vectors which are used to produce compositions of the
CC invention.
XX
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 NTYRLKLYSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTFDFEGHSINDYISPD 60
Db 51 NTYRLKLYSLRWISDHEYLKQENNLVFNFAEYGNSSVFLENSTFDFEGHSINDYISPD 110
Qy 61 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 120
Db 111 GQFILLEYNVVKQWRHSYTASYDIYDLNKRQLITEERIPNNTQWVTWSPVGHKLAYVWNN 170
Qy 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAQFND 180
Db 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAQFND 230
Qy 181 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDLSQYPKTVRVPYPKAGAVNPTVKFFVNTDSLSVNTNATSIQITAP 290
Qy 241 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQERISLQWLRRIQNYSVMDICDYDESSGRWNCCLVARQHIEMST 350
Qy 301 TGWVGRFRPSEPHTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFTITKGTWEVIGIEAL 360
Db 351 TGWVGRFRPSEPHTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFTITKGTWEVIGIEAL 410
Qy 361 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQOL 420
Db 411 TSDYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQOL 470
Qy 421 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVMPSKKLDFFILNETKFWYQML 480
Db 471 RCGSGPLPLYTLTHSSVNDKGLRVLEDNSALDKMLQNVMPSKKLDFFILNETKFWYQML 530
Qy 481 PPHFDKSKKYPILLDVIYAGCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 540
Db 531 PPHFDKSKKYPILLDVIYAGCSQKADTVFRLNWTATYLASTENIIIVASFDFGRSGYQGDKI 590
Qy 541 MHAINRRLGTFFVEDQIEAARQFSKMGFVDNKRRIALWGSYGGYVTSWVLGSGGVFKCG 600

Db 591 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSMLVSGSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSVYTERYMGFLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVYTERYMGFLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIRKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIRKQCFSLP 766

RESULT 13
AD071644
ID ADO71644 standard; protein; 766 AA.
XX
AC ADO71644;
XX
DT 26-AUG-2004 (first entry)
XX

XX Amino acid sequence of a human CD26 protein.
DE CD26; chemotherapeutic; radiotherapeutic; cancer; cell growth;
KW dipeptidyl peptidase IV; DPPIV; topoisomerase II inhibitor;
KW cell cycle arrest; tumour; tumour necrosis; immune response; human.
XX
OS Homo sapiens.
XX
XX WO2004045497-A2.
PN
XX
PD 03-JUN-2004.
XX
XX 15-MAY-2003; 2003WO-US015499.
XX
XX 17-MAY-2002; 2002US-0381606P.
PR
XX (TEXA) UNIV TEXAS SYSTEM.
PA
XX Dang NH, Morimoto C;
XX
XX WPI; 2004-420511/39.
DR N-PSDB; ADO71643.
DR
XX
XX Use of a CD26 composition, and a chemotherapeutic and/or a
PT radiotherapeutic agent for e.g. inhibiting the cell growth, inducing cell
PT cycle arrest; killing a cancer cell, treating cancer, or inducing tumor
PT regression or tumor necrosis.
XX

PS Claim 23; Page 175-176; 182pp; English.
XX
XX The specification describes a CD26 composition which, in conjunction with
XX chemotherapeutic or radiotherapeutic agents, is used for the treatment
XX of cancer. Expression of CD26 enhances the sensitivity of
XX the cancer cell to the chemotherapeutic or radiotherapeutic agent. CD26
XX is a dipeptidyl peptidase IV (DPPIV). The chemotherapeutic agent is a
XX topoisomerase II inhibitor. The CD26 composition of the invention is
XX useful for inhibiting the growth of a cell, inducing cell cycle arrest in
XX a cell, killing a cancer cell, potentiating the effect of a
XX chemotherapeutic agent and/or a radiotherapeutic agent on a tumour cell,
XX inducing or enhancing apoptosis of a cancer cell, treating cancer, or
XX inducing tumour regression or tumour necrosis. The CD26 composition is
XX further useful for increasing topoisomerase II expression in a cell, for
XX activating an antigen-presenting cell, or for potentiating immune
XX responses of an animal. The present sequence represents a CD26 protein,
XX and is encoded by vectors which are used to produce compositions of the
XX invention.
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 NTYRLKLYSLRWISDHELYLKQENNLVFNAEYGNSSVFLNSTDFFDGHISINDYSISPD 60
Db 51 NTYRLKLYSLRWISDHELYLKQENNLVFNAEYGNSSVFLNSTDFFDGHISINDYSISPD 110
Qy 61 GQFILLLENNYVVKQWRHSYTSASYDIYDLNKRQLITEERI PNTQWTVTSPVGHKLAYWNN 120
Db 111 GQFILLLENNYVVKQWRHSYTSASYDIYDLNKRQLITEERI PNTQWTVTSPVGHKLAYWNN 170
Qy 121 DIVVKLEPNLPYSRIITWTGKEDIYNGITDWTVEEVEFSAYSAALWMSPNGTFLAYAOFPND 180
Db 171 DIVVKLEPNLPYSRIITWTGKEDIYNGITDWTVEEVEFSAYSAALWMSPNGTFLAYAOFPND 230
Qy 181 TEVPLIEYSFYSDSLQYPKTVRPYPKAGAVNPTVKFFVYVNTDLSLSSVTNATSIQITAP 240
Db 231 TEVPLIEYSFYSDSLQYPKTVRPYPKAGAVNPTVKFFVYVNTDLSLSSVTNATSIQITAP 290
Qy 241 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYVMDICDYDESSGRWNCVLARQHIEMST 300
Db 291 ASMLIGDHYLCDVTWATQERISLOWLRRIQNYVMDICDYDESSGRWNCVLARQHIEMST 350
Qy 301 TGWVGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKKDCTFITKGTWVIGIEAL 360
Db 351 TGWVGRFRPSEPHFTLDGNSFYKIIISNEGYRHCYFQIDKKDCTFITKGTWVIGIEAL 410
Qy 361 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
Db 411 TSDLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
Qy 421 RCGSGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFTILNETKFWYQML 480
Db 471 RCGSGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFTILNETKFWYQML 530
Qy 481 PPHFDKSKYPLLLDVYAGPCOKADTVFRLNWTATYLASTENIIVASFDCRGSYGQGDKI 540
Db 531 PPHFDKSKYPLLLDVYAGPCOKADTVFRLNWTATYLASTENIIVASFDCRGSYGQGDKI 590
Qy 541 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSMLVSGSGVFKCG 600
Db 591 MHAINRLGTFEVEDQIEAARQFSKMGFVDNKRITAIWGSYGGYVTSMLVSGSGVFKCG 650
Qy 601 IAVAPVSRWEYDVSVYTERYMGFLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
Db 651 IAVAPVSRWEYDVSVYTERYMGFLPTPDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
Qy 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIRKQCFSLP 716
Db 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIRKQCFSLP 766

RESULT 14

ABM80355
ID ABM80355 standard; protein; 766 AA.

AC ABM80355;

DT 18-NOV-2004 (first entry)

XX Tumour-associated antigenic target (TAT) polypeptide PRO80881, SEQ:895.
XX
XX Tumour-associated antigenic target; TAT; human; overexpression; cancer;
KW tumour; diagnosis; cell proliferative disorder; breast cancer;
KW colorectal cancer; lung cancer; ovarian cancer; liver cancer;
KW central nervous system cancer; bladder cancer; pancreatic cancer;
KW cervical cancer; melanoma; leukaemia; hybridisation probe;
KW chromosome identification; chromosome mapping; gene mapping;
XX gene therapy; cytostatic.

OS Homo sapiens.

XX WO2004030615-A2.

XX 15-APR-2004.

XX

PP 29-SEP-2003; 2003WO-US028547.
XX
PR 02-OCT-2002; 2002US-0414971P.
XX
PA (GETH) GENENTECH INC.
XX
PI Wu TD, Zhang Z, Zhou Y;
XX
DR WPI; 2004-347921/32.
DR N-PSDB; ACN37783.
XX
New tumor-associated antigenic target polypeptides and nucleic acids,
PT useful in preparing a medicament for treating or detecting a
PT proliferative disorder, e.g. breast, lung, colorectal, ovarian or
PT prostate cancer or tumor.
XX
PS Claim 12; SEQ ID NO 895; 7273pp; English.
XX
CC The invention relates to human tumour-associated antigenic target (TAT)
CC polypeptides, and their related nucleic acids. The TAT polypeptides are
CC overexpressed in cancer tissues compared to normal tissues, and may thus
CC serve as effective targets for the diagnosis and treatment of cancer in
CC mammals. The invention also relates to nucleic acid and polypeptide
CC sequences at least 80% identical to the TAT nucleic acids and
CC polypeptides; expression vectors and host cells comprising a TAT nucleic
CC acid; an antibody specific for a TAT polypeptide; a peptide or organic
CC molecule which binds to a TAT polypeptide; fusion proteins comprising a
CC TAT polypeptide; and methods and compositions for the treatment or
CC diagnosis of cancer in mammals. TAT polypeptides, nucleic acids,
CC antibodies, antagonists, binding molecules and compositions are useful
CC for diagnosing or treating a cell proliferative disorder associated with
CC increased TAT expression, particularly cancers such as breast cancer,
CC colorectal cancer, lung cancer, ovarian cancer, liver cancer, bladder
CC cancer, pancreatic cancer, cervical cancer, cancers of the central
CC nervous system, melanoma and leukaemia. TAT nucleic acids may further be
CC used as hybridisation probes, in chromosome and gene mapping, in
CC chromosome identification and in gene therapy. The present sequence
CC represents a TAT polypeptide of the invention
XX
SQ Sequence 766 AA;

Query Match 100.0%; Score 3877; DB 8; Length 766;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 NTVRLKLYSLRWISDHELYLKQENNLVFNAYGNSVFLNSTFDFGHSINDYSISP 60
DB 51 NTVRLKLYSLRWISDHELYLKQENNLVFNAYGNSVFLNSTFDFGHSINDYSISP 110
QY 61 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWNN 120
DB 111 GQFILLEYNVYKQWRHSYTASYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWNN 170
QY 121 DIVVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSPNGTFLAYAQFND 180
DB 171 DIVVKIEPNLPSYRITWTGKEDIYNGITDWTVEEVEFSAYSALWSPNGTFLAYAQFND 230
QY 181 TEVPLIEYSFYSDLESQPKTVRPVYPKAGAVNPTVKFFVYNTDSLSSVTNATSIQITAP 240
DB 231 TEVPLIEYSFYSDLESQPKTVRPVYPKAGAVNPTVKFFVYNTDSLSSVTNATSIQITAP 290
QY 241 ASMLIGHYLCVDTWATQERISLOWLRRIQNYSMVMDICDYDESSGRWNCVLVARQHIE MST 300
DB 291 ASMLIGHYLCVDTWATQERISLOWLRRIQNYSMVMDICDYDESSGRWNCVLVARQHIE MST 350
QY 301 TGMVGRFRPSPHFTLDGNSFYKII SNEEGRHICYFQIDKDKCTFTTKGTWEVIGIEAL 360
DB 351 TGMVGRFRPSPHFTLDGNSFYKII SNEEGRHICYFQIDKDKCTFTTKGTWEVIGIEAL 410
QY 361 TSDLYLYISNEYKMGPMGGRNLYKIQLSDYTKVTCLSCELNPERCQYVSFSKEAKYYQL 420
DB 411 TSDLYLYISNEYKMGPMGGRNLYKIQLSDYTKVTCLSCELNPERCQYVSFSKEAKYYQL 470

QY 421 RCGSGPLPLYTLHSSVNDKGLRLVLEDNLSALDKMLQNVQMPSSKKLDFFIILNETKFWQMIL 480
DB 471 RCGSGPLPLYTLHSSVNDKGLRLVLEDNLSALDKMLQNVQMPSSKKLDFFIILNETKFWQMIL 530
QY 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDRGSGYQGDKI 540
DB 531 PPHFDKSKYPLLLDVYAGPCSKADTVFRLNWTATYLASTENIIIVASFDRGSGYQGDKI 590
QY 541 MHAINRLGTFFVEDQIEAARQFSKMGFVDNKRRIALIWGWSYGGYVTSWVLGSGSGVFKCG 600
DB 591 MHAINRLGTFFVEDQIEAARQFSKMGFVDNKRRIALIWGWSYGGYVTSWVLGSGSGVFKCG 650
QY 601 IAVAPVSRWEYVDSVYTERYMGILPTPEDNLNHYRNVSTVMSRAENFKQVEVLLIHGTADDN 660
DB 651 IAVAPVSRWEYVDSVYTERYMGILPTPEDNLNHYRNVSTVMSRAENFKQVEVLLIHGTADDN 710
QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 716
DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFHKQCFSLP 766
RESULT 15
ADP54458
ID ADP54458 standard; protein; 766 AA.
XX ADP54458;
XX
DT 18-NOV-2004 (first entry)
DE Human PRO protein sequence SEQ ID NO:434.
XX human; PRO; immune related disease; inflammatory immune response;
KW immune response stimulation; anti-allergic; antianaemic; antiarthritic;
KW antirheumatic; antidiabetic; antiinflammatory; antipsoriatic;
KW antirheumatic; antithyroid; CNS; dermatological; gastrointestinal;
KW haemostatic; hepatotropic; immunostimulant; immunosuppressive; muscular;
KW nephrotropic; neuroprotective; osteopathic; respiratory; vasotropic;
KW virucide; gene therapy.
XX
OS Homo sapiens.
XX
PN WO2004039956-A2.
XX
PD 13-MAY-2004.
XX
PF 28-OCT-2003; 2003WO-US034381.
XX
PR 29-OCT-2002; 2002US-0422472P.
XX
PA (GETH) GENENTECH INC.
XX
PI Aggarwal S, Clark H, Gurney AL, Schoenfeld J, Williams PM;
PI Wood WI, Wu TD;
XX
XX WPI: 2004-376182/35.
DR N-PSDB; ADP54457.
XX
PT New PRO polynucleotides and polypeptides, useful in useful in diagnosing
PT and treating an immune related disease, e.g. systemic lupus
PT erythematosus, rheumatoid arthritis, diabetes mellitus or asthma and in
PT stimulating an immune response.
XX
XX Claim 1; SEQ ID NO 434; 3009pp; English.
PS
CC The present invention describes an isolated PRO nucleic acid (1). Also
CC described: (1) a vector comprising (1); (2) a host cell comprising the
CC vector of (1); (3) a process for producing a PRO polypeptides; (4) an
CC isolated PRO polypeptide; (5) a chimeric molecule comprising the
CC polypeptide of (4) fused to a heterologous amino acid sequence; (6) an
CC antibody which specifically binds to a polypeptide of (4); (7) a
CC composition of matter comprising a polypeptide of (4), an agonist or
CC antagonist of the polypeptide or an antibody that binds to the
CC polypeptide in combination with a carrier; (8) an article of manufacture

Search completed: January 30, 2006, 14:22:35
Job time : 138 secs

comprising a container, a label on the container and a composition of matter of (7); (9) a method of treating an immune related disease in a mammal; (10) a method for determining the presence of a PRO polypeptide in a sample suspected of having the polypeptide; (11) a method of diagnosing an immune related disease or an inflammatory immune response in mammal; (12) a method of identifying a compound that inhibits or mimics the activity of or expression of a gene encoding a PRO polypeptide; and (13) a method of stimulating the immune response in a mammal. The PRO sequences have anti-allergic, antianaemic, antiarthritic, antiaesthetic, antidiabetic, anti-inflammatory, antipruritic, antirheumatic, antithyroid, CNS, dermatological, gastrointestinal, haemostatic, hepatotropic, immunostimulant, immunosuppressive, muscular, nephrotropic, neuroprotective, osteopathic, respiratory, vasotropic and virucide activities, and can be used in gene therapy. The nucleic acid (I) and the encoded polypeptides, compositions, kits and methods are useful in diagnosing and treating an immune related disease and in stimulating an immune response. The present sequence represents a human PRO protein from the present invention.

XX Sequence 766 AA;

Query Match	100.0%;	Score 3877;	DB 8;	Length 766;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 716;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	NTYRLKLSLRWISDHEVLYKQENNILVFNAEYGNSSVFLNENSTDFEFGHSINDYSISPD	60	
Db	51	NTYRLKLSLRWISDHEVLYKQENNILVFNAEYGNSSVFLNENSTDFEFGHSINDYSISPD	110	
Qy	61	GQFILLBYNKKWRHSYTSYDIYDLNKKQLITEERIPNNTQWVTWSPVGHKLAYVWNN	120	
Db	111	GQFILLBYNKKWRHSYTSYDIYDLNKKQLITEERIPNNTQWVTWSPVGHKLAYVWNN	170	
Qy	121	DIYVKIEPNLPSRITWTGKEDIITNGITDWVVEEVFSAYSALWSPNGTFLAYAQFND	180	
Db	171	DIYVKIEPNLPSRITWTGKEDIITNGITDWVVEEVFSAYSALWSPNGTFLAYAQFND	230	
Qy	181	TEVPLIEYSFSDLESLOYPKTVRPYPKAGAVNPTKFFVVTNDSLSSVTNATSIQITAP	240	
Db	231	TEVPLIEYSFSDLESLOYPKTVRPYPKAGAVNPTKFFVVTNDSLSSVTNATSIQITAP	290	
Qy	241	ASMLIGHYLCVDTWATQERISLOWLRRIQNSVMDICDYDESSGRWNCCLVARQHBMST	300	
Db	291	ASMLIGHYLCVDTWATQERISLOWLRRIQNSVMDICDYDESSGRWNCCLVARQHBMST	350	
Qy	301	TGWGFRPSPHPHTLDGNSFYKIIISNEEGYRHCYQIDKQCTFITKGTWEVIGIEAL	360	
Db	351	TGWGFRPSPHPHTLDGNSFYKIIISNEEGYRHCYQIDKQCTFITKGTWEVIGIEAL	410	
Qy	361	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL	420	
Db	411	TSDYLYYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAKYYQL	470	
Qy	421	RCSGPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFTILNETKFWQMIL	480	
Db	471	RCSGPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFTILNETKFWQMIL	530	
Qy	481	PPHFDKSKKYPPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	540	
Db	531	PPHFDKSKKYPPLLDVYAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQGDKI	590	
Qy	541	MHAINRLRGTFEVEDQIEAARQFSKMGFVDNKRITATWGSYGGVVTSMVLGSGGVFKCG	600	
Db	591	MHAINRLRGTFEVEDQIEAARQFSKMGFVDNKRITATWGSYGGVVTSMVLGSGGVFKCG	650	
Qy	601	IAPVPSRWEYSDSVYTERYWGMLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	660	
Db	651	IAPVPSRWEYSDSVYTERYWGMLPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	710	
Qy	661	VHFQQAQISKALVDGVDFQAMWYTDHGHASSTAHOHIYTHMSHFIKQCFSLP	716	
Db	711	VHFQQAQISKALVDGVDFQAMWYTDHGHASSTAHOHIYTHMSHFIKQCFSLP	766	

THIS PAGE BLANK (USP 16)

GenCore version 5.1.6
Copyright (c) 1993 - 2006 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:14:31 ; Search time 41 Seconds
(without alignments)
1680.273 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877
Sequence: 1 NTVRLKLYSLRWISDHEYL.....AHQHITHSHFIKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3871	99.8	766	1 CDHU26	dipeptidyl-peptida
2	3353.5	86.5	792	1 A39914	dipeptidyl-peptida
3	3336	86.0	760	1 S23752	dipeptidyl-peptida
4	1930.5	49.8	759	2 I38593	fibroblast activat
5	1117	28.8	803	2 I68600	dipeptidyl aminope
6	1117	28.8	865	2 I54331	dipeptidyl aminope
7	1094	28.2	803	2 A41793	dipeptidyl aminope
8	940	24.2	793	2 T41703	dipeptidyl aminope
9	908	23.4	818	1 A30107	dipeptidyl aminope
10	791.5	20.4	711	2 S66261	X-Pro dipeptidyl-p
11	745	19.2	829	2 T19514	hypothetical prote
12	726	18.7	931	2 A49737	dipeptidyl aminope
13	720	18.6	779	2 T25173	hypothetical prote
14	716	18.5	799	2 T25174	hypothetical prote
15	628	16.2	738	2 A87516	dipeptidyl-peptida
16	594.5	15.3	741	2 JC5142	X-Pro dipeptidyl-p
17	590.5	15.2	743	2 T37700	probable dipeptidyl
18	445.5	11.5	931	2 T32919	hypothetical prote
19	332	8.6	795	2 F82858	dipeptidyl-peptida
20	272	7.0	657	2 E70025	probable acylamino
21	244.5	6.3	764	2 JC8016	acylaminoacyl-pept
22	222.5	5.7	622	2 F71174	hypothetical prote
23	221	5.7	591	2 H72474	probable acylamino
24	219	5.6	642	2 C71137	hypothetical prote
25	206.5	5.3	632	2 E75057	peptidase PAB1418
26	205	5.3	631	2 H75007	probable acylamino
27	205	5.3	683	2 E87495	prolyl oligopeptid
28	191	4.9	709	2 B82580	alanyl dipeptidyl
29	188	4.8	659	2 F72568	probable acylamino

probable acylamino
prolyl oligopeptid
acylamino-acid-rel
conserved hypothet
prolyl oligopeptid
F44B9.1 protein -
acylaminoacyl-pept
hypothetical prote
probable peptide h
probable acylamino
probable acylamino
hypothetical prote
acylaminoacyl-pept
oligopeptidase B (
acylaminoacyl-pept
hypothetical prote

ALIGNMENTS

RESULT 1

CDHU26

dipeptidyl-peptidase IV (EC 3.4.14.5) - human

N/Alternate names: cell surface glycoprotein CD26; thymocyte-activating molecule (THAM)

C/Species: Homo sapiens (man)

C/Date: 31-Dec-1993 #sequence revision 23-Aug-1996 #text change 09-Jul-2004

C/Accession: S24313, B42408, A42408, B61136, S59510, I56154, S59857, S15520

R/Misumi, Y.; Hayashi, Y.; Arakawa, F.; Ikeharu, Y.

Biochim. Biophys. Acta 1131, 333-336, 1992

A/Title: Molecular cloning and sequence analysis of human dipeptidyl peptidase IV, a se

A/Reference number: S24313, MUID:92329551, PMID:1352704

A/Accession: S24313

A/Molecule type: mRNA

A/Residues: 1-6, 'I', '8-766 <MIS>

A/Cross-references: UNIPROT:P27487; UNIPARC:UPI000016AE80; EMBL:X60708; NID:G35335; PID:

R/Barbault, D.; Lacasa, M.; Baricault, L.; Marguet, D.; Sapin, C.; Trotot, P.; Barbat, A.

J. Biol. Chem. 267, 4824-4833, 1992

A/Title: Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like colon cance

IV mRNA levels during cell differentiation.

A/Reference number: A42408, MUID:92165847, PMID:1347043

A/Accession: B42408

A/Molecule type: mRNA

A/Residues: 1-5, 'R', '7-436, 'S', '438-556, 'I', '558-662, 'E', '664-766 <DAR1>

A/Cross-references: UNIPARC:UPI0000052ACB; GB:M80536; NID:G181569; PIDN:AAA52308.1; PID

A/Experimental source: intestine

A/Note: this sequence corresponds with the author's translation

A/Accession: A42408

A/Molecule type: mRNA

A/Residues: 1-5, 'R', '7-436, 'S', '438-556, 'I', '558-662, 'E', '664-711, 'G', '713-766 <DAR2>

A/Cross-references: UNIPARC:UPI0000172A2B; GB:M80536; NID:G181569

A/Note: sequence extracted from NCBI backbone (NCBI:83986, NCBI:83988); this sequence

R/Gorvel, J.P.; Ferrero, A.; Chambraud, L.; Rigal, A.; Bonicel, J.; Maroux, S.

Gastroenterology 101, 618-625, 1991

A/Title: Expression of sucrose-isomaltase and dipeptidylpeptidase IV in human small int

A/Reference number: A61136; MUID:91317403; PMID:1677636

A/Accession: B61136

A/Molecule type: protein

A/Residues: 1-15, 'X', '17-22 <GOR>

A/Cross-references: UNIPARC:UPI0000172A2C

R/Boehm, S.K.; Gum Jr., J.R.; Erickson, R.H.; Hicks, J.W.; Kim, Y.S.

Biochem. J. 311, 835-843, 1995

A/Title: Human dipeptidyl peptidase IV gene promoter: tissue-specific regulation from a

A/Reference number: S59510; MUID:96067599; PMID:7487939

A/Accession: S59510

A/Molecule type: DNA

A/Residues: 1-31 <BOE>

A/Cross-references: UNIPARC:UPI000016B4A6; GB:S79876; NID:G1195574; PIDN:AAB35614.1; PI

R/Tanaka, T.; Camerini, D.; Seed, B.; Torimoto, Y.; Dang, N.H.; Kameoka, J.; Dahlberg,

J. Immunol. 149, 481-486, 1992

A/Title: Cloning and functional expression of the T cell activation antigen CD26.

A/Reference number: I56154; MUID:92325476; PMID:1352530

A/Accession: I56154

QY 477 QMILPPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQ 536
 Db 547 QILKPATFTDTTHYPLLLVDGTPGSGVAEKFEVSWETVMVSHGAVVVKCDGRSGGFQ 606
 QY 537 GDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIALIAGWSYGGYVTSWVL---GSG 593
 Db 607 GTKLLHEVRRLGLLEKQDMEAVRTMLKEQYIDRTFVAVFGKDYGGYLSTYILPAKGEN 666
 QY 594 SG-VFKCGIAPVPSRWEYVDSVYTERYMGLPPTPEDNLDH--YRNSTVMSRAENFKQVEY 650
 Db 667 QCGTFTCGSALSPTIDPKLYASAFSERYLGL----HGLDNRAVEMTKVAHRVSALBEQQF 722
 QY 651 LLIHGTADDNVHQSSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHQHIYTHMSHF 710
 Db 723 LIHPTADEKIHQHTAELITQLIRKANYSLQIYPDESHYFTSSSLKQHLYSIINFFV 782
 QY 711 QCFSL 715
 Db 783 ECFRI 787

RESULT 6

I54331
 dipeptidyl aminopeptidase like protein - human
 C:Species: Homo sapiens (man)
 C:Date: 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change 09-Jul-2004
 C:Accession: I54331
 R:Yokotani, N.; Doi, K.; Wentholt, R.J.; Wada, K.
 Hum. Mol. Genet. 2, 1037-1039, 1993
 A:Title: Non-conservation of a catalytic residue in a dipeptidyl aminopeptidase IV-related
 A:Reference number: I54331; MUID:93372805; PMID:8103397
 A:Accession: I54331
 A:Status: preliminary;
 A:Molecule type: mRNA
 A:Residues: 1-865 <RES>
 A:Cross-references: UNIPROT:P42659; UNIPARC:UPI0000047378; GB:M96859; NID:g306705; PIDN:
 C:Superfamily: dipeptidyl-peptidase IV

Query Match 28.8%; Score 1117; DB 2; Length 865;
 Best Local Similarity 33.2%; Pred. No. 3.3e-69;
 Matches 241; Conservative 150; Mismatches 290; Indels 44; Gaps 20;

QY 11 RWISDHEYLK-QENNILVFNAYGNSSVPLENSTDFEGHSINDYSISPDGQFILLEYN 69
 Db 149 KWISDTEFYREQGTVRLNVNVTNTSTVLEGGKIESL-RAIR-YEISPDREYALFSYN 206
 QY 70 YVQWRHSYTSYDIYDL--NKRQLITEERIPN-NTQWTVTSPVGHKLAVVWNDIYVKI 126
 Db 207 VEPIYQHSYTYGYVLSKIPHGDPQSLDPPPEVSNALQYAGWGPKGQQLIFIPENNIYCA 266
 QY 127 EPNLPSYRITWTKGEDIYNGITDWYEEVEFSAYSALMWSPNGTFLAYAOENDETVPLI 186
 Db 267 HVCKQAIRVSVTSGEGVIYNGLSDWLYEEILKTHIAHWSPDGTRLAYAINDSRVPIM 326
 QY 187 EYSFYSDESLQYPKTVRVPYKAGAVNPVKFFVNTDSLSVNTNATSIQTAPASMLIG 246
 Db 327 ELPTYTGSII--YPTKVPYHPYKAGSENPISLHVI---GLNGPTH--DLEMPDPDPRMR 379
 QY 247 DHLCDVTWATQERISLOWLRRIQNVSMDCIDYDESSGRWNCLVARQHIEMSTTCGWVGR 306
 Db 380 EYITWVKWATSKVAVTWLNRAQNSILTLG--DATTG---VCTKKH-EDSEAWLHR 432
 QY 307 FRPSEPHFTLDGNSFYKIIISNEEGRYHCYF-----QIDKKDCTFTTKGTWVEIGIE 358
 Db 433 -QNEEPVFSKDGKRFKFFIRAIPOGGGKFFHITVSSSQPNSSNDNIQTSITSGDWDVTKIL 491
 QY 359 AL--TSDLYYIINEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAK 416
 Db 492 AYDEKGNKIYFLSTE--DLPRRQLYSANTEGNNFNQCLSCDL- VENTCYFVSASFHSD 548
 QY 417 YYQLRSGPGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLLFIILNETKFWY 476
 Db 549 FFLKCEGEGVPMVTVHNTTDDKKMFDELNETNEHVKAINDRQMPKVEYRDIEIDYNNLPM 608

QY 477 QMILPPHFDKSKYPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIIVASFDGRSGGYQ 536
 Db 609 QILKPATFTDTTHYPLLLVDGTPGSGVAEKFEVSWETVMVSHGAVVVKCDGRSGGFQ 668
 QY 537 GDKIMHAINRRLGTFEVEDQIEAARQFSKMGFVDNKRIALIAGWSYGGYVTSWVL---GSG 593
 Db 669 GTKLLHEVRRLGLLEKQDMEAVRTMLKEQYIDRTFVAVFGKDYGGYLSTYILPAKGEN 728
 QY 594 SG-VFKCGIAPVPSRWEYVDSVYTERYMGLPPTPEDNLDH--YRNSTVMSRAENFKQVEY 650
 Db 729 QCGTFTCGSALSPTIDPKLYASAFSERYLGL----HGLDNRAVEMTKVAHRVSALBEQQF 784
 QY 651 LLIHGTADDNVHQSSAQISKALVDVGVDFQAMWYTTDEHGIASSTAHQHIYTHMSHF 710
 Db 785 LIHPTADEKIHQHTAELITQLIRKANYSLQIYPDESHYFTSSSLKQHLYSIINFFV 844
 QY 711 QCFSL 715
 Db 845 ECFRI 849

RESULT 7

A41793
 dipeptidyl aminopeptidase-like protein (EC 3.4.14.-) short form, DPPX-S - bovine
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 04-Mar-1993 #sequence_revision 18-Nov-1994 #text_change 09-Jul-2004
 C:Accession: A41793
 R:Wada, K.; Yokotani, N.; Hunter, C.; Doi, K.; Wentholt, R.J.; Shimasaki, S.
 Proc. Natl. Acad. Sci. U.S.A. 89, 197-201, 1992
 A:Title: Differential expression of two distinct forms of mRNA encoding members of a di-
 A:Reference number: A41793; MUID:92108018; PMID:1729689
 A:Accession: A41793
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-803 <WAD>
 A:Cross-references: UNIPROT:P42659; UNIPARC:UPI000002A83C; GB:M76429; NID:g408719; PIDN:
 C:Superfamily: dipeptidyl-peptidase IV
 C:Key words: dipeptidylpeptide hydrolase; glycoprotein
 F:57,342/Binding site: carbohydrate (Aen)

Query Match 28.2%; Score 1094; DB 2; Length 803;
 Best Local Similarity 32.9%; Pred. No. 1.2e-67;
 Matches 239; Conservative 153; Mismatches 287; Indels 48; Gaps 22;

QY 11 RWISDHEYLK-QENNILVFNAYGNSSVPLENSTDFEGHSINDYSISPDGQFILLEYN 69
 Db 87 KWISDTEFYREQGTVRLNVNVTNTSTVLEGGKIESL-RAIR-YEISPDREYALFSYN 144
 QY 70 YVQWRHSYTSYDIYDL--NKRQLITEERIPN-NTQWTVTSPVGHKLAVVWNDIYVKI 126
 Db 145 VEPIYQHSYTYGYVLSKIPHGDPQSLDPPPEVSNALQYAGWGPKGQQLIFIPENNIYCA 204
 QY 127 EPNLPSYRITWTKGEDIYNGITDWYEEVEFSAYSALMWSPNGTFLAYAOENDETVPLI 186
 Db 205 HVCKQAIRVSVTSGEGVIYNGLSDWLYEEILKTHIAHWSPDGTRLAYAINDSRVPIM 264
 QY 187 EYSFYSDESLQYPKTVRVPYKAGAVNPVKFFVNTDSLSVNTNATSIQTAPASMLIG 246
 Db 265 ELPTYTGSII--YPTKVPYHPYKAGSENPISLHVI---GLNGPTH--DLEMPDPDPRMR 317
 QY 247 DHLCDVTWATQERISLOWLRRIQNVSMDCIDYDESSGRWNCLVARQHIEMSTTCGWVGR 306
 Db 318 EYITWVKWATSKVAVTWLNRAQNSILTLG--DATTG---VCTKKH-EDSEAWLHR 370
 QY 307 FRPSEPHFTLDGNSFYKIIISNEEG---YRHICYF---QIDKKDCTFTTKGTWVEIGIE 358
 Db 371 -QNEEPVFSKDGKRFKFFIRAIPOGGGKFFHITVSSSQPNSSNDNIQTSITSGDWDVTKI- 428
 QY 359 ALTSD-----YLYYIINEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKE 414
 Db 429 -LSYDEKRSQIYFLSTE--DLPRRQLYSASTVSGSPNQCLSCDL- VDNCTYFVSASFSPG 484

Db 136 NGQNLVESITASPLKRLKRLIRTSVQNRHSTFGSVFVVDKSSSF--EEIGNEVALA 192
Qy 106 TSPVGHKLAYVMNDIYVIEPNLPSYRITWTGKEDIINGIDWVYEEVFSAYSALW 165
Db 193 IWSPNNDIAVQDNNIYISAISSKTIKRAVNDGSSFLNGKFDWYEEVEFDDKAAW 252
Qy 166 WSPNGTFLAQAOFNDTEVP--LIBYSYDESLOYKTVRPVYPKAGAVNPTVKFFVNT 223
Db 253 WSPGDLAFLKIDSESGEFLIIPY-YVQDEKDIYPEMRISIKYPSKTPNPHABLWY-- 309
Qy 224 DLSSTVNATSIQITAPASMLIGHYLCVDTWATQERISLQWLRRIQNSVMDICDYDES 283
Db 310 ----SMKDCSTSPRISGNKDGSLLTETVWVGNGVNLVKTDRSSDILITFLDITIAK 365
Qy 284 SGRWNCILVARQHIMSTTGM-----VGRFRPSEPHFTLDGNSFYKIIISNEEGYRHICYFQ 338
Db 366 TSN-----VVRN--ESSNGMWETHNTLFIIPANETFORPHGVYDILP-IGGYNHLAYFE 418
Qy 339 -IDKDCFTFKGTWEVI-GIEALTS--DYLYISNEYKMGPGGRNLYKIQL----SDYTK 391
Db 419 NSNSSHYKTLTEGKWEVNGVPLAFDSMENRLYFISTRKSTE--RHVYIIDLRSPNEIE 476
Qy 392 VTCLSCELNPERCOYYSVSFSKAKYVQLRCSGGLP---LYTLHS-----SYNDK 439
Db 477 VTDTSB-----GVYDVSFSSGRFRGLLTVKGPVYQKIVDPHRSKAKCKDKGNVLGK 530
Qy 440 GLRVLENSALDKMLQNVQMPKSKLDPIILNETKFWYQM-----ILPPHFK--SKYP 491
Db 531 SLYHLEKNEVLTKILEYAVPRKSFRELNLGKDEFGKDLVNSYEILPNDFDELSDHP 590
Qy 492 LLDVYAGPCSKADTVFRLNWTATYLASTENIIVASPDGRSGYQGDKIMHAINRRLGTF 551
Db 591 VFFPAYGPNSSQVVKTFVSFGFNEVWASQLNAIVVVDGRGTGKGDPRSLVRDLGDI 650
Qy 552 EVEDOIEAARQFSKMGVDNKRITAIWGSYGGYVTSVLGSGSG-VFKGCIAPVSRWE 610
Db 651 EARDQISAAISLYGLTTFVDPKISLFGSYGGYVTLTKLEKGGRHFPKYGMSVAPVTDWR 710
Qy 611 YDVSVYTERYMGLTPTEDNLDHY-----RNSVTMSRAENFKQVEYLLIHGTAADNVHFQ 665
Db 711 FYDSVYTERY--HTQENFDGVSVSVHNTALQANRF-----LLMHGTGDDNVHFQ 763
Qy 666 SAQISKALVDVGDV-FOAMWYTTDEDHGASSTAHQHIYTHMSHFIKQCF 713
Db 764 SLKFLDLLDLNGVENVDVHVPFSDHSIRYHNVANVIVFDKLDWAKRAF 812

RESULT 10
S66261
X-Pro dipeptidyl-peptidase (EC 3.4.14.11) - Flavobacterium meningosepticum
C;Species: Flavobacterium meningosepticum
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 09-Jul-2004
C;Accession: S66261
R;Kobashima, T.; Yoshida, T.; Ito, K.; Yoshimoto, T.
Arch. Biochem. Biophys. 320, 123-128, 1995
A;Title: Cloning, sequencing, and expression of the dipeptidyl peptidase IV gene from *F.*
A;Reference number: S66261; MUID:95314307; PMID:7793970
A;Accession: S66261
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-711 <KAB>
A;Cross-references: UNIPROT:Q47900; UNIPARC:UPI00000B2115; EMBL:D42121; NID:g577283; PID
C;Superfamily: dipeptidyl-peptidase IV
C;Keywords: dipeptidyl-peptidase hydrolase

Query Match 20.4%; Score 791.5; DB 2; Length 711;
Best Local Similarity 30.3%; Pred. No. 8.6e-47;
Matches 216; Conservative 125; Mismatches 274; Indels 97; Gaps 24;

Qy 31 AEYGNSSVFLNSTFDFEGHSHINDYSISPDGQFILLLEYVYVVKWRHSYTSYDIYDLNKR 90
Db 60 AKYSYTSQKERNIVD---GSFGYTFNSDESKILLQKSSQSIYRHSFLGKFEVKDLKSR 116

Qy 91 QLITEBIPNNTQW---TWSPVGHKLAYVMNDIYVIEPNLPSYRITWTGKEDIYNG 147
Db 117 TVVS-----LNNANWIEPKFSPDGSKVAFIADNNLFLQDLNTKTIQITITDGKKEIING 172
Qy 148 ITDWVVEEVEFSAYSALWSPNGTFLAQAOFNDTEVPLEIBYSYDESLOYKTVRPVYP 207
Db 173 LGDWVYEE--FGHADYVQWKNAGDALVFRFDRKVPINPIIYQ--LYPKLMTYKYP 229
Qy 208 KAGAVNPTVKFFVNTDLSSTVNATSIQITAPASMLIGHYLCVDTWAT--QERISLOW 265
Db 230 KAGEENSAN-----TAYLYQLSSGSAQLNFSS-----EKYIIPQLFQTNANDEIVAT 279
Qy 266 LRRIQNSVMDICD-----YDESSGRWNCILVARQHIMSTTGMVGRFRPSPHF 314
Db 280 ANRHQN--KYDLLKVTAKTAASKLFTETDNW---IETDNLME----- 319
Qy 315 TLDGNSFYKIIISNEEGYRHICYFQIDKDCFTFKGTWEVIGIEALTSYLYISNEYK 374
Db 320 FLDDNSFL-WASERDGRHLYWYDAAGKLKQVSKGDWEII-----NYYGY 364
Qy 375 MPGGRNLYKIQLSDYTKVTCSCEL--NPERCOY-----SVSFSKAKYVQLRCSGP 425
Db 365 NPATKEVY-IQTTEKSINKVSKLINTGKTQLLSNAEGNNSAASKTNYFINTGSTA 423
Qy 426 GLPLYTLHSSVNDKGLRVLENSALDKMLQNVQMPKSKLDPIIL-----NETKFWYQMIL 480
Db 424 KVPTKYLKANGKDVKEIQNDDLLKLSDNFIK--EFITIPNAAGQMNW--MIK 479
Qy 481 PPHFDKSKYPLLLDVYAGPCSKADTVFRLN---WATYLASTENIIVASPDGRSGYQG 537
Db 480 PKNFDPAKKYPVFMFQYSGGQVANSWDGGNGIWFDMLAQ-KGYLVVCDVGRGTGFRG 538
Qy 538 DKIMHAINRRLGTFEVEDOIEAARQFSKMGVDNKRITAIWGSYGGYVTSVLGSGGVF 597
Db 539 TTKYKVTYKNGLKYETEDQITAAKWLGNOSYVDKSRIGIFGYSYGGYMASLAMTKGADV 598
Qy 598 KGIAPVAPVSRWEYDVSVYTERYMGLTPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 657
Db 599 KMGIAVAPVNTMRFYDSIYTERF--LOTQENKDGVDLNSPTTYAKLLKG-KELLIHGTA 655
Qy 658 DNVHFPQQAQSKALVDVGVDFQAMWYTTDEDHGASSTAHQHIYTHMSHFI 709
Db 656 DNVHFPQNSMEFSEALIQNKQDFDMAYDPDKNHSIIGGNTRPOLYKEMTNYI 707

RESULT 11
T19514
Hypothetical protein C27C12.7 - *Caenorhabditis elegans*
C;Species: *Caenorhabditis elegans*
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T19514
R;Thomas, K.
submitted to the EMBL Data Library, March 1996
A;Reference number: Z19134
A;Accession: T19514
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-829 <WIL>
A;Cross-references: UNIPROT:Q18253; UNIPARC:UPI000007DCS3; EMBL:Z69883; PIDN:CAA93743.1
A;Experimental source: clone C27C12
C;Genetics:
A;Gene: CBSP:C27C12.7
A;Map position: X
A;Introns: 4/3; 51/1; 123/3; 166/3; 188/2; 279/2; 392/3; 501/3; 553/1; 583/1; 606/2; 64
C;Superfamily: dipeptidyl-peptidase IV

Query Match 19.2%; Score 745; DB 2; Length 829;
Best Local Similarity 28.8%; Pred. No. 1.8e-43;
Matches 227; Conservative 128; Mismatches 282; Indels 150; Gaps 35;

Qy 8 YSLRWISDHEYLKQENNILVFNAEYGNSSVFLNSTFDFEGHSHINDYSISPDGQFILL 65
Db 105 YDIWLPDGSFVQMDD----FTIRKMKKIPLGSSVAEPFFN-----NGEVVKAL 151

```
QY 66 ---LEYNVYKQ-----WRHSYTSYDIYDLANKQLITEERI PNTTQW-----V 105
Db 152 SSMKAYGSKKYNELWRHSAEYLIVHIVKNNKTSTVE-----QHWGPEENSIIQAF 204
QY 106 TWSP--VGHKLAYVNNNDIYVKLEPNLPSVRIWT--GKEDIINGITDWYEEVEFSAYS 162
Db 205 YNPNASSNDFVYVHYNLYYQDKPEKPGAIQLTVGGSTFRFGLANWLYEEIELEASS 264
QY 163 ALWMSRNGTFLAYAQNDTEVPLIEYSFYSDLESLOPKTVRPYPKAGAVNPT--VKFVV 221
Db 265 AVWSPSGRVSYLRFDDREVNRIFLPKYTD--SYVEYFELPYKAGVQNNLTVQYIW 323
QY 222 NTDSLSVYTNATSIQITAPASMLIGHYLCVDTWAT-----QERISLOWLRIONY 272
Db 324 DSENHKIVETAPPNELSAAN-----GDYYLTNKNWITMPRNGSDLGBERLVTWANDQNH 379
QY 273 SVMNDICDYDESSGRWNCIVA-----ROHIEMSTTGMWGRFRPSE-----PHTILD 317
Db 380 VYFSLCNEQ-----CVMALSTQFSDNRQLWVSPKDVGVF--PTETGFLTVPKHKDD 432
QY 318 GNSFYKIISNEGYRHICYFQIDKDCFTFITKGTWE-----VIG-----IEALT----- 361
Db 433 GNI-----YNHVAHVELDGTGTGKITWIGENFEDVILVGYSSKIDALTFSAYGD 482
QY 362 --SDLYIYISNEYKMGPGGRNLYKIQLSDYTKTCLSCELNPERCQY--SVSFSKEAYY 418
Db 483 GVGEFSTYIVRE-----AMYSNKKTTLLQKVT-----QPBDCKTLGQSADPTGQRI 529
QY 419 QLRCSGP--GLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDFIIL--NETKEW 475
Db 530 VVOCEKPFONTRILYV--DVDDTKIMLEGTT--KAVIPFDVFNKFKGLKLPDSIGDH 585
QY 476 YQMLPFPHPDKSKYPLLLDVYAGPCS-----QKADTVFRLNWTATYLASTENIIVASFQGR 531
Db 586 YNMLTPANLLDGAKIPLLLDIYGPDSKQVFKTPTAH---AIQIVSQYDIAYARIDVR 641
QY 532 GSGYQGDKMIAHNRRLGTFEVDQIEAARQF--SKMGFVDNKRRIATWGSYGYVTSMWL 590
Db 642 GTGGRGMDVKEAYVRKLGDAEVVDTLDMIRAFINTFTGFIDEDRIAVMGWSYGGFLTSKIA 701
QY 591 GSGSG--VFKGIAVAPVSRWEYVDSVTRYMGLPTPEDNLDHRYNSTWMSRAENPKQVE 649
Db 702 IKDQGLVKCAISIAPTVDFKYYSAYTERYLQ--QPAENLQYIINTVIPHARNTVTK 759
QY 650 YLLIHGTADDNVHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHQ--HIYTHMSH 707
Db 760 YLLAHGERDDNVHYQNSARWSEALQQNGIHFTQLVYANEAH-----SLSHKLFHLYGEVQR 815
QY 708 FI-KQCF 713
Db 816 FLMNDCF 822

RESULT 12
A49737
dipeptidyl aminopeptidase (EC 3.4.14.-) - yeast (Saccharomyces cerevisiae)
N;Alternate names: protein O5045; protein YOR219c; protein YOR50-9
C;Species: Saccharomyces cerevisiae
C;Date: 03-Mar-1994 #sequence revision 03-Mar-1994 #text_change 09-Jul-2004
C;Accession: A49737; S45451; S60946; S67112; S71721
R;Santa Anna-A, S.; Herskowitz, I.
A;Reference number: A49737
submitted to the Protein Sequence Database, July 1993
A;Molecule type: DNA
A;Residues: 1-931 <SAN>
A;Cross-references: UNIPROT:P33894; UNIPARC:UPI0000136060; GB:L21944; NID:g347196; PIDN:
R;Anna-Arriola, S.S.; Herskowitz, I.
Yeast 10, 801-810, 1994
A;Title: Isolation and DNA sequence of the STE13 gene encoding dipeptidyl aminopeptidase
A;Reference number: S45451; MUID:95066382; PMID:7975997
A;Accession: S45451
```

```
A;Molecule type: DNA
A;Residues: 1-931 <ANN>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:L21944; NID:g347196; PIDN:AAA35119.1;
R;Gallison, F.; Dujon, B.
submitted to the EMBL Data Library, October 1995
A;Description: Sequence and analysis of a 33 kb fragment from the right arm of chromosome
A;Reference number: S60938
A;Accession: S60946
A;Molecule type: DNA
A;Residues: 1-931 <GAL>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
R;Boyer, J.; Fairhead, C.; Gaillon, L.; Gallison, F.; Michaux, G.; Thierry, A.; Dujon, B.
submitted to the Protein Sequence Database, July 1996
A;Reference number: S67104
A;Accession: S67112
A;Molecule type: DNA
A;Residues: 1-931 <BOY>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:Z75127; NID:g1420507; PIDN:CAA99437.1;
R;Gallison, F.; Dujon, B.
Yeast 12, 877-885, 1996
A;Title: Sequence and analysis of a 33 kb fragment from the right arm of chromosome XV
A;Reference number: S71713; MUID:96437977; PMID:8840505
A;Accession: S71721
A;Status: nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-931 <GAM>
A;Cross-references: UNIPARC:UPI0000136060; EMBL:X92441; NID:g1050762; PIDN:CAA63182.1;
A;Note: the nucleotide sequence was submitted to the EMBL Data Library, October 1995
C;Genetics:
A;Gene: SGD:STE13; YC11
A;Cross-references: SGD:S0005745; MIPS:YOR219c
A;Map position: 15R
C;Function:
A;Description: involved in processing of alpha-factor prepropheromone
C;Keywords: dipeptidylpeptide hydrolase; glycoprotein; transmembrane protein
F;113-150/Domain: transmembrane #status predicted <TM>
F;377/Binding site: carbohydrate (Aen) (covalent) #status predicted
F;785,863,896/Active site: Ser, Asp, His #status predicted

Query Match 18.7%; Score 726; DB 2; Length 931;
Best Local Similarity 26.9%; Pred. No. 4.4e-42;
Matches 195; Conservative 146; Mismatches 318; Indels 66; Gaps 23;

QY 29 FNAEYGNSSVFLNENSTFDFGHSINDYSISPGQFILLEYNVYKQMRHYSYASYDIYDLN 88
Db 223 FEWNLGGRFLYEGVEFTVSTVQIN-YKLDK----LIFGTNLESPRHSKGFYWKLDN 277
QY 89 K--RQLITEERIPNNTQ-----WVTWSPVGHKLAYVNNNDIYVKEPNLPSYRITWT 138
Db 278 TGNIEPILPPEKSDNDNYELGSLSYAHFSPAYNTIYFVYNNLFLQQVNSGVAKKVTED 337
QY 139 GKEDIYNGITDWYEEVEFSAYSALWMSRNGTFLAYAQNDTEVPLIEYSFYSDLESQY 198
Db 338 GSKD-IFNAKPDWIIYEEVSLASDQAIWAPDDSKAVFARFNDTSVDDIRLNRYTNWNEAY 396
QY 199 PKTVRVPYKAGAVNPTVKFFVVNTDSLSVYTNATSIQITAPASMLIGHYLCVDTWATQ 258
Db 397 LSDTKIKYKPKGFQNPQFDLFLVN-----LQNGIYSINTGGQK---DSILYNGKWISP 447
QY 259 ERISLOWLRIONYQVSMIDICYDESSGRWNCIVARQHITEMSTGTWGRFR-----PSEPH 313
Db 448 DTFRFEITDR--NSKILDVKVYDIPSSQ--MLTVRNTNSNLFNGWIEKTKDILSIPPKPE 503
QY 314 FTLDGNSFYKIISNEEGYRHICYF--QIDKDCFTFITKGTWEV-----IGIEALTSDLYY 367
Db 504 LKRMGYGIDIHADSRGSHLFYFTVFAKEPIQLTKGNWETGNGIVGEYET--DTIFF 562
QY 368 ISNEYKMGPGGRNLYKIQLSDYTKTCLSCELNP--ERCQYYSVFSKEAKYVQLRCSGPG 426
Db 563 TANEIGVM--SQHLYSISLSTTTQNTFQSLQNPSPDKYDFYFELSSSARYAISKKLQPD 620
QY 427 LPYLTILH-----SSVNDKGLRVLEDNSALDKMLQNVOMPSKKLDFIILNE-TKFWYQ 477
```

Db 621 TPIKAVAGPLTRVLNVAEIHDDSIQLTKDEKFKIKNYDLPTSYKTMVLDDGVEINYI 680
QY 478 MILPPHFDKSKYPLLLDVTAGPCSKADTVFRLNWTATYLASTENIIVASFDRGSGYQG 537
Db 681 EIKPANLNPKKKYPILNIIYGGPGSQFTTKSSLAFAQAVVSGLDVTLVQIEPRGTGGK 740
QY 538 DKIMHAINRRRLGTPEVEDQIEAARQFQKMG--FVDNKRIAIWGWSYGYVTSWVLGSGG 595
Db 741 WSPRSWAREKLYGWEPRDITVTKKFIQRNSQIHDESKIAIWGWSYGGFTSLKTVELDNG 800
QY 596 -VFKCGIAPVPSRWEYDYSYTERYMGFLPTPEDNLDHYNSTVMSRAENFKQVEYL--- 651
Db 801 DTFKYAMAVAPVTNTWLYDSYTERYMNQPS-ENHEGYFEVSTI---QNFKSPESLKL 855
QY 652 -LTHGTADDNVHFOQSAISKALVDGV--DFOAMWYTDDEHGIIASSTAHOHYTHSHFI 709
Db 856 FIVHGTDDNVHIOHTFRVLQDLNLLGLTNYDMHIFPDSHDSIRYHNAQRIVFKLYYL 915
QY 710 KQCFS 714
Db 916 RDAFA 920

RESULT 13

T25173
hypothetical protein T23F1.7a - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25173
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25173
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: bna
A;Residues: 1-779 <WIL>
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI000002A220; EMBL:Z81129; PIDN: CAB03411.1
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CRSP:T23F1.7a
A;Map position: 5
A;Introns: 13/3; 52/1; 111/3; 151/2; 177/3; 346/3; 504/1; 537/1; 628/2; 683/1; 723/1
C;Superfamily: dipeptidyl-peptidase IV

Query Match 18.6%; Score 720; DB 2; Length 779;
Best Local Similarity 27.6%; Pred. No. 8.8e-42;
Matches 218; Conservative 124; Mismatches 296; Indels 152; Gaps 33;

QY 2 TYRLKLYSLRWISDHELYLKYQ---ENNILVFNA-----EYGNSSVFLNSTFDFEGH 50
Db 60 SWRLNVSDLRSL--RYPYAEFAFTDNNAVMQSWEGVEIVDGVSRLLI-----FGR 108
QY 51 SINDYSISPDGQ---FILLFYN-----YVQWRHSYASYDYIDLNKRQLITE 95
Db 109 E-NGAETPSADRKRYFAMMDHAPNPGWNPQNETFHLKIVNNNERITYDI-GLAKEESV-- 164
QY 96 ERIPNNTQWTVSPVGHKLAVYWNNDIYVKIEBNLPSYRITWTGKEDIIVNGITDWWYEE 155
Db 165 -----IQAFKNGKGFNDFFVESNKIYQSSPEEGLTRVNSGGEHTV--DGLFDWYEE 217
QY 156 EVFSAYSALWSPNGTFLAYAQFNDTEVPLIEYSFYS-----DESLOYPKTVRPYPK 208
Db 218 EIFGRKDMWMSSTGQDLAYASDNHLLTKVNSLKTTHRLPEYPIDTNFHPKPT----- 271
QY 209 AGAVNPVKPFVYNTDLSLSTVNTATSIQITAPASMLIGDHYLQDVTW-----ATQERISLQ 264
Db 272 -AKVLPY-----TLSIWNKTEQSRQLDVQLKDSLSYHLLAVKWLINGTEQLVSV- 323
QY 265 WLRRIQNSYWDICDYDESSGRNCLVARQHIE--MSTTGWGRFRPSEPHFTLDGNSFY 322
Db 324 WTRNYQNEVALTICD-----WDTAICRLEFEYKYKASKRWVTH-----DDFH 364

QY 323 KIISNEBGRHICVFOI--DKDKCT-----PITGKTWEVIGIEALTS 362
Db 365 SITSPED-----TLFFLLPHDRDNAPQOAVASRLSHGQLRTPKFLNIGEDVDTVSINGINK 420
QY 363 DYLYIYSNEYKMGPGGRNLKYIQLSDYTK--VTCLSCELNPERCQYYSVSFSKEAKYYQL 420
Db 421 ETRTIFPHAAAPKPSHRSLFSYSLADESRNSAYCISCSI--KNCTWAAQAMDDQMKTAIV 478
QY 421 RCGSGPLPLYLTHSSV-----NDKGLRVLEDNALSADKMLQNVQMPSKKLFIILNFKF- 474
Db 479 SCKGPAAP-----HTAIVNLTRMDSDDKKEHANLLYDKTYQN--RVEEAGLPVLIKETIKIS 533
QY 475 -----WYQMLPHPF---DKSKYKPYLLDYYAGPCSKADTVFRLNWTATYLASTENIIVA 526
Db 534 DDFDALIKLSIPKDIYNRDKHQAIPLIHVHYGPNQNTKEATQIGIEEVVASASQAAIL 593
QY 527 SFGRSGVGQDQKIMHAINRRRLGTFEVEDQIEAARQFSKM--GFVDNKRIAIWGWSYGY 584
Db 594 RIDGRSGGRGWKYSALYQGLGTFEVEDQIKAVVLRLLYRHLDDARRVAVFGWSYGGF 653
QY 585 VT--SMVLGSGGVFKCGIAPVPSRWEYDYSYTERYMGFLPTPEDNLDHYNSTVMSRAE 643
Db 654 MTLSMYNEAPEQFFKCAVSVAPVTNFAYYDATYTERYMG---DAPLESY--SDVTKKLD 707
QY 644 NFKQVEYLLIHGTADDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIIASSTAHOHYT 703
Db 708 NFKSTRLLMHGLLDDNVHFOQSAISKALVDGVDFQAMWYTDDEHGIIASSTAHOHYT 703
QY 704 HMSHFIFKQCF 713
Db 766 KMTFLRLQCF 775

RESULT 14

T25174
hypothetical protein T23F1.7b - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25174
R;Wilkinson, J.
submitted to the EMBL Data Library, October 1996
A;Reference number: Z19990
A;Accession: T25174
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-799 <WIL>
A;Cross-references: UNIPROT:O18119; UNIPARC:UPI0000061190; EMBL:Z81129; PIDN: CAB03412.1
A;Experimental source: clone T23F1
C;Genetics:
A;Gene: CRSP:T23F1.7b
A;Map position: 5
A;Introns: 13/3; 52/1; 111/3; 197/3; 366/3; 524/1; 557/1; 648/2; 703/1; 743/1
C;Superfamily: dipeptidyl-peptidase IV

Query Match 18.5%; Score 716; DB 2; Length 799;
Best Local Similarity 27.5%; Pred. No. 1.7e-41;
Matches 221; Conservative 122; Mismatches 300; Indels 162; Gaps 33;

QY 2 TYRLKLYSLRWISDHELYLKYQ---ENNI-----LVFNABYQ----- 34
Db 60 SWRLNVSDLRSL--RYPYAEFAFTDNNAVMQSWEGVEIVDGVSRLLIFGRENGABITP 116
QY 35 -----NSSVFLNSTFDFEGHSINDYSISPDGQQLILEYNYVVKWHSYTA 80
Db 117 SADRKYFAMMDHAPNPGWNPQNETF--HLKIVNNNERLNP-----LLPFE--VEELFRELSD 169
QY 81 SYDIYDLNKRQLITEERIPNNTQWTVSPVGHKLAVYWNNDIYVKIEBNLPSYRITWTGK 140
Db 170 SRITYDIGLRK---EESV---IQAFKNGKGFNDFFVESNKIYQSSPEEGLTRVNSGG 223
QY 141 EDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAQFNDTEVPLIEYSFYS-----D 193
Db 224 EHTV--DGLFDWYEEEIFGRKDMWMSSTGQDLAYASDNHLLTKVNSLKTTHRLPEYPID 282

```
QY 194 ESLOYKTRVPYPKAGAVNPTVKFFVWNTDSLSSVTNATSIQITAPASMLIGHYLCV 253
Db 283 TNHYPKTF-----AKVLPTY-----TLSIWNKKTQESRQLDQVQLKDSLYHYLLAV 329
QY 254 TW-----ATQBRISLQWLRRRQNTQSVMDICDYDESSGRWNCCLVARQHIE--MSTTGWVGRF 307
Db 330 KWEINGTEQLVSV-WTNRVQNEVALTICD-----WDAICRLFEFYKYASKRWVTH- 380
QY 308 RPSEPHFTLDGNSFYKIIISNEEGYRHCYFOI--DKKDCI-----FI 347
Db 381 -----DDFHSITSPEI-----TLFLLPHDKRDNAFQVVASLRLSHGQLRTPKFL 425
QY 348 TKGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTK--VTCLSCELNPERCQ 405
Db 426 NLGEYDVTSINGINKETRTIFFHAAAPKPSHRSLFSYSLADESRNSAYCISCSI--KNCT 483
QY 406 YYSVSFSKEAKYYQLRCSGPGPLPLYTLHSSV-----NDKGLRVLNDSALDKMLQNVQMP 460
Db 484 WAQAQWDDQMTAIVCSCKGPAAP-----HTAIVNLTRMDSDDKTEHANLLYDKTYQN-RVE 538
QY 461 SKKLDFFIILNETKP-----WYQMLPPHF-----DKSKKYPLLLDVYAGPCSQKADTVFRL 511
Db 539 EAGLPVIKETIKISDDFDALIKLSIPKOIYNRDKHQAIPLIVHVYGGPNDQNTKEATOI 598
QY 512 NWATYLASTENIIVASFDRGSGYOGDKIMHAINRRLTGTEVEVDQIEAARQFSKM--GFV 569
Db 599 GIEEVVASAQAAILRIDRGRSGGRGMKYSALYQGLTGTEVEVDQIKAVLRLRYRHLL 658
QY 570 DNKRIAWGWSYGYVT-SMVLGSGGVFKGIAVAPSVRWEYYSVYTERYVGLPTPED 628
Db 659 DARRVAVGWSYGGFWTLNWNREAPQFFKCAVSAVPNTNFAYDATYTERYVNG-----DA 714
QY 629 NLPHYRNSTVMSRAENFKQVEYLLIHGTADDNVHFOQSAQISKALVDVGVDFQAMWYTD 688
Db 715 PLESY--SDVTCKLDNFKSTRLLMHGLLDNDNVHFNQNSAILIDELQNRGVDFDLVWVPNQ 772
QY 689 DHGIASTAHQHIYTHMSHFIKQCF 713
Db 773 AHSLSRST--HVVGKMTFLRQCF 795

RESULT 15
A87516
dipeptidyl peptidase IV [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 09-Jul-2004
C:Accession: A87516
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: A87516
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-738 <STO>
A:Cross-references: UNIPROT:Q9A6E0; UNIPARC:UPI00000C7616; GB:AE005673; NID:g13423647; H
C:Genetics:
A:Gene: CC2154

Query Match 16.2%; Score 628; DB 2; Length 738;
Best Local Similarity 26.8%; Pred. No. 1.9e-35;
Matches 189; Conservative 115; Mismatches 269; Indels 132; Gaps 23;

QY 52 INDYSISPDGQFILLVYN---YVQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTWS 108
Db 111 IVEYSWDRQGRFLVPLDGLYL-----DAVAGKITRLTE--TPGDEVDAKVS 157
QY 109 PVGHKLAYVMNDIYVKIEPNLFSYRITWTGKEDIYINGITDWVYEEVFSAYSALWSP 168
Db 158 PKGGYVSYVRDQNLXIKPVAGGAETALTDTGKDALSF-GVAEFIVQEEL-DREFTGYWSP 215
```

```
QY 169 NGTFLAYAQPNDTEVPLIEYSFYSDESLQYPKTVRVYPKAGAVNPTVKFFVWNTDSLSS 228
Db 216 DESRIYVTRVDESQVDIVP---RADTGGGATVMNQRYPRAGRPNAVVDLF----- 263
QY 229 VTNATSIQITAPASMLIGHYLCVTVW-ATQBRISLQWLRRRQNTQSVMDICDYDESSGRW 287
Db 264 VDDLASGKVTAIDLGNKNDIYVARVNASADGKTVVYVORLSRDQ--KTLDLLAFDAATGAG 331
QY 288 NCLVARQHITEMTGTGWGRFRPSEPHFTLDGNSFYKII-----SNEEGYRHCYFQID 340
Db 322 KTILT-----DTDPHFIEVSNDRPLTDTGTFLWSEKDGNOHLRYAAD 365
QY 341 KDCTFITKGTWEVIGIEALTSYLYIISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELN 400
Db 366 GKLIQAQITKGDNDPVTGLEG-----VDEARKVAIFASID 399
QY 401 -PERCQYYSVSFSKEAKYYQLRCSG-----PGLPLYTLHSSVNDK 439
Db 400 TPIERLYEVSYAKPKPKALTSAGGWAAKVADNGGAFAGTYSDPKTSPQATLYSADGK 459
QY 440 GLRVLEDNSALD-----KMLQNVQMP-----KKLDFIILNETKFWYQMLPPHFDKSKKY 490
Db 460 RVRWIEENKLAEGHPWPVYAANLPQEPGSLKAAD---GET-LHYEILKPIGDFPAKKY 514
QY 491 PLLLDVYAGPCSQKADTVFRLW-----ATYLASTENIIVASFDRGSGYOGDKIMHAIN 545
Db 515 PAIVSVYGGPHQARV-----MKWHSFSERTYLEA--GYVIFKLDNRSGNSAKFWRALD 568
QY 546 RLRLTGTEVEVDQIEAARQFSKMGVFNKRIAIWGSYGGVYVTSVVLGSGGVFKGIAVAP 605
Db 569 RKLGTVEVEDQLLGAFLASQYVVDADKLGVMGWSYGGFWALMLLTAEVTPFKAGAGAP 628
QY 606 VSRWEYYSVYTERYVGLPTPEDNLDRNNTVMSRAENFKQVEYLLIHGTADDNVHFOQ 665
Db 629 PTWMSLYDTAYTERYVNG--KPDENKAGYAYSINNRRIDKLAPGSLLLHLHGMDNDNVIFEN 686
QY 666 SAQISKALVDVGVDFQAMWYTDDEHGIASSTAH--QHIYTHMSHF 708
Db 687 STRLMAALQKRAILFEMAMYGERHSAPGSKTKGLSVLKTHLDDFF 731
```

Search completed: January 30, 2006, 14:23:27
Job time : 44 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2006, 23:58:33 ; Search time 11 Seconds
(without alignments)

704.874 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTYRLKLYSLRWISDHELYL.....AHQHYTHMSHFHKQCSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 75621 seqs, 10829074 residues

Total number of hits satisfying chosen parameters: 75621

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA New:

- 1: /cgn2_6/ptodata/2/pubppaa/US08_NEW_PUB pep.*
- 2: /cgn2_6/ptodata/2/pubppaa/US06_NEW_PUB pep.*
- 3: /cgn2_6/ptodata/2/pubppaa/US07_NEW_PUB pep.*
- 4: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep.*
- 5: /cgn2_6/ptodata/2/pubppaa/US09_NEW_PUB pep.*
- 6: /cgn2_6/ptodata/2/pubppaa/US10_NEW_PUB pep.*
- 7: /cgn2_6/ptodata/2/pubppaa/US11_NEW_PUB pep.*
- 8: /cgn2_6/ptodata/2/pubppaa/US60_NEW_PUB pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	6	US-10-522-789-2
2	3834	98.9	762	7	US-11-116-939-13
3	2138	55.1	760	7	US-11-186-284-55
4	525	13.5	882	7	US-11-151-601-20
5	272	7.0	657	7	US-11-179-977-1
6	124	3.2	668	6	US-10-454-437-118
7	115	3.0	1593	6	US-10-453-372-1092
8	112.5	2.9	1436	6	US-10-453-372-1094
9	110	2.8	19	7	US-11-116-939-27
10	110	2.8	72	7	US-11-151-601-22
11	109.5	2.8	877	7	US-11-077-550-157
12	109	2.8	2516	6	US-10-647-956A-2
13	107	2.8	866	7	US-11-077-550-32
14	107	2.8	871	7	US-11-077-550-155
15	107	2.8	873	7	US-11-077-550-163
16	107	2.8	873	7	US-11-077-550-165
17	107	2.8	876	7	US-11-077-550-66
18	107	2.8	878	7	US-11-077-550-62
19	107	2.8	879	7	US-11-077-550-30
20	107	2.8	908	7	US-11-077-550-64
21	107	2.8	914	7	US-11-077-550-60
22	107	2.8	949	7	US-11-077-550-68
23	107	2.8	1130	7	US-11-077-550-44
24	107	2.8	1132	7	US-11-077-550-46
25	106	2.7	878	7	US-11-077-550-12

26	106	2.7	907	7	US-11-077-550-16	Sequence 16, Appl
27	106	2.7	953	7	US-11-077-550-14	Sequence 14, Appl
28	106	2.7	1013	7	US-11-077-550-18	Sequence 18, Appl
29	106	2.7	4495	6	US-10-453-372-1002	Sequence 1002, Ap
30	105.5	2.7	871	7	US-11-077-550-2	Sequence 2, Appli
31	105.5	2.7	871	7	US-11-077-550-8	Sequence 8, Appli
32	105.5	2.7	871	7	US-11-077-550-26	Sequence 26, Appl
33	105.5	2.7	871	7	US-11-077-550-153	Sequence 153, App
34	105.5	2.7	873	7	US-11-077-550-6	Sequence 6, Appli
35	105.5	2.7	873	7	US-11-077-550-149	Sequence 149, App
36	105.5	2.7	873	7	US-11-077-550-151	Sequence 151, App
37	105.5	2.7	873	7	US-11-077-550-167	Sequence 167, App
38	105.5	2.7	873	7	US-11-077-550-169	Sequence 169, App
39	105.5	2.7	879	7	US-11-077-550-159	Sequence 159, App
40	105.5	2.7	894	7	US-11-077-550-4	Sequence 4, Appli
41	105.5	2.7	1127	7	US-11-077-550-40	Sequence 40, Appl
42	105.5	2.7	1129	7	US-11-077-550-42	Sequence 42, Appl
43	105.5	2.7	3003	6	US-10-453-372-1080	Sequence 1080, Ap
44	105.5	2.7	3361	6	US-10-453-372-1082	Sequence 1082, Ap
45	104.5	2.7	658	6	US-10-873-528-17	Sequence 17, Appl

ALIGNMENTS

RESULT 1

US-10-522-789-2

; Sequence 2, Application US/10522789

; Publication No. US20050260732A1

; GENERAL INFORMATION:

; APPLICANT: TANABE SEIYAKU CO., LTD.

; TITLE OF INVENTION: Three-dimensional structure of dipeptidyl peptidase IV

; FILE REFERENCE: 03-039-PCT

; CURRENT APPLICATION NUMBER: US/10/522,789

; CURRENT FILING DATE: 2005-01-28

; PRIOR APPLICATION NUMBER: US 60/398,761

; PRIOR FILING DATE: 2002-07-29

; NUMBER OF SEQ ID NOS: 2

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 2

; LENGTH: 766

; TYPE: PRT

; ORGANISM: Homo sapiens

; ORGANISM: Homo sapiens

US-10-522-789-2

Query Match	100.0%;	Score 3877;	DB 6;	Length 766;
Best Local Similarity	100.0%;	Pred. No. 2.3e-299;		
Matches	716;	Conservative	0;	Mismatches 0;
		Indels	0;	Gaps 0;
Qy	1	NTYRLKLYSLRWISDHELYLKOENNLVFNNAEYGNSSVFLNENSTFDFGHSINDYSISPD	60	
Db	51	NTYRLKLYSLRWISDHELYLKOENNLVFNNAEYGNSSVFLNENSTFDFGHSINDYSISPD	110	
Qy	61	GQFILLEYNVKKWRHSYASVDYDLNKRQLITEIRIPNNTQMTWSPVGHKLAYVWNN	120	
Db	111	GQFILLEYNVKKWRHSYASVDYDLNKRQLITEIRIPNNTQMTWSPVGHKLAYVWNN	170	
Qy	121	DIYVKLEPNLPSYRIITWTKGEDIYNGITDWTVEEVFSAYLSALWSPNGTFLAYAQFND	180	
Db	171	DIYVKLEPNLPSYRIITWTKGEDIYNGITDWTVEEVFSAYLSALWSPNGTFLAYAQFND	230	
Qy	181	TEVPLEYGFYSDSELYQPKTVRVPYKAGAVNPTVKFFVWNTDSLSSVTNATSIQTAP	240	
Db	231	TEVPLEYGFYSDSELYQPKTVRVPYKAGAVNPTVKFFVWNTDSLSSVTNATSIQTAP	290	
Qy	241	ASMLIGDHLYCDVTWATQERISLQWLRRIQNTYSVMDICDYDESSGRNCLVARQHIEMST	300	
Db	291	ASMLIGDHLYCDVTWATQERISLQWLRRIQNTYSVMDICDYDESSGRNCLVARQHIEMST	350	
Qy	301	TGWVGRFRSEPHFTLDGNSFYKLIISNEGYRHCYFQIDKKDCTTITKGTWEVIGIEAL	360	
Db	351	TGWVGRFRSEPHFTLDGNSFYKLIISNEGYRHCYFQIDKKDCTTITKGTWEVIGIEAL	410	

QY 361 TSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQYL 420
 DB 411 TSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQYL 470
 QY 421 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWQMIL 480
 DB 471 RCGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFWQMIL 530
 QY 481 PPHFDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 540
 DB 531 PPHFDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGYQGDKI 590
 QY 541 MHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSWVLGSGGVFKCG 600
 DB 591 MHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSWVLGSGGVFKCG 650
 QY 601 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 660
 DB 651 IAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHGTADDN 710
 QY 661 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 716
 DB 711 VHFQSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSLP 766

RESULT 2
 US-11-116-939-13
 ; Sequence 13, Application US/11116939
 ; Publication No. US2005026595A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Stephen Tomlinson
 ; APPLICANT: Richard J. Quig
 ; FILE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
 ; CURRENT FILING DATE: 2005-04-28
 ; PRIOR FILING DATE: 2005-04-28
 ; NUMBER OF SEQ ID NOS: 27
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 13
 ; LENGTH: 762
 ; TYPE: PRT
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
 ; OTHER INFORMATION: construct
 US-11-116-939-13

Query Match 98.9%; Score 3834; DB 7; Length 762;
 Best Local Similarity 98.6%; Pred. No. 5.8e-296;
 Matches 711; Conservative 0; Mismatches 6; Indels 6; Gaps 1;
 QY 2 TYRLKLY-----SLRWISDHELYKQENNLVFNAYEYGNSSVFNLENSTDFEFGHSINDY 55
 DB 42 TYTLTDLYKNYKLLRWISDHELYKQENNLVFNAYEYGNSSVFNLENSTDFEFGHSINDY 101
 QY 56 SISPDGQFILLEYYNKKQMRHSYASYDIYDLNKRQLITEERIPNNNTQWTVWSPVGHKLA 115
 DB 102 SISPDGQFILLEYYNKKQMRHSYASYDIYDLNKRQLITEERIPNNNTQWTVWSPVGHKLA 161
 QY 116 YVWNNDIYKIEBPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWWSPNGTFLAY 175
 DB 162 YVWNNDIYKIEBPNLPSYRITWTGKEDIYNGITDWMYEEVFSAYSALWWSPNGTFLAY 221
 QY 176 AQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPVKFVNTDSLSSVTNATSI 235
 DB 222 AQFNDTEVPLIEYSFYSDESLOYPKTVRVPYKAGAVNPVKFVNTDSLSSVTNATSI 281
 QY 236 QITAPASMLIGDHYLCDDVTWATERISLOWLRRIQYNSVMDICDYDESSGRMNCNVARQH 295
 DB 282 QITAPASMLIGDHYLCDDVTWATERISLOWLRRIQYNSVMDICDYDESSGRMNCNVARQH 341

QY 296 IEWSTTGWGRPSPPEHFTLDGNSFYKLIISNEEGYRHCYFQIDKCKOCTFTKGTWEVI 355
 DB 342 IEWSTTGWGRPSPPEHFTLDGNSFYKLIISNEEGYRHCYFQIDKCKOCTFTKGTWEVI 401
 QY 356 GIEALTSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEA 415
 DB 402 GIEALTSDYLYYSNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEA 461
 QY 416 KYQRLRCSGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFW 475
 DB 462 KYQRLRCSGPGGLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDFFILNETKFW 521
 QY 476 YQMLPPHFDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGY 535
 DB 522 YQMLPPHFDKSKYPLLLDVVAGPCSKADTVFRLNWTATYLASTENIIVASFDGRGSGY 581
 QY 536 QGDKIMHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSWVLGSGSG 595
 DB 582 QGDKIMHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGGVYVTSWVLGSGSG 641
 QY 596 VFKCGIAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHG 655
 DB 642 VFKCGIAVAPVSRWEYDVSYYTERYMGFLPTPEDNLDHYRNSVTMSRAENFKQVEYLLIHG 701
 QY 656 TADDNVHFOOSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSL 715
 DB 702 TADDNVHFOOSAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSL 761
 QY 716 P 716
 DB 762 P 762

RESULT 3
 US-11-186-284-55
 ; Sequence 55, Application US/11186284
 ; Publication No. US20050266493A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Millennium Pharmaceuticals, Inc.
 ; APPLICANT: Berger, Allison
 ; APPLICANT: Guillemette, Tracy L.
 ; APPLICANT: Kamatkar, Shubhangi
 ; APPLICANT: Schlegel, Robert
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Thibodeau, Stephen N.
 ; APPLICANT: BURGART, Lawrence J.
 ; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
 ; TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
 ; TITLE OF INVENTION: THERAPY OF COLON CANCER
 ; FILE REFERENCE: MP001-029P2RNM
 ; CURRENT APPLICATION NUMBER: US/11/186,284
 ; CURRENT FILING DATE: 2005-07-21
 ; PRIOR APPLICATION NUMBER: US/10/301,822
 ; PRIOR FILING DATE: 2002-11-21
 ; PRIOR APPLICATION NUMBER: US 60/339,971
 ; PRIOR FILING DATE: 2001-12-10
 ; PRIOR APPLICATION NUMBER: US 60/361,978
 ; PRIOR FILING DATE: 2002-03-05
 ; PRIOR APPLICATION NUMBER: US 60/381,988
 ; PRIOR FILING DATE: 2002-05-20
 ; NUMBER OF SEQ ID NOS: 228
 ; SOFTWARE: FASTSEQ for Windows Version 4.0
 ; SEQ ID NO 55
 ; LENGTH: 760
 ; TYPE: PRT
 ; ORGANISM: Homo Sapiens
 US-11-186-284-55

Query Match 55.1%; Score 2138; DB 7; Length 760;
 Best Local Similarity 53.3%; Pred. No. 1.3e-161;
 Matches 383; Conservative 134; Mismatches 187; Indels 14; Gaps 7;
 QY 2 TYRLKLYSLRWISDHELYKQ-ENNILVFNAYEYGNSSVFNLENSTDFEFGHSIN--DYSIS 58

```
Db 51 TFSYKTEFPNWSQOEVLHOSADNNVILYNIETQSYTILSNRTW-----KSVNASNYGLS 106
Qy 59 PDGOFILLEVNVKOWRHSYASDYIDVLANRQLITTEERIPNNTQWTTWSPVGHKLAYVW 118
Db 107 PDROFVLESYKLMRYSTATYIIYDLSNGEFTVRGNELPRPIQYLCWSPVSGKLAYVY 166
Qy 119 NNDIYKIEPLNPSYRITWTGKEDIYNGITDWDYVEEVFSAYSALWSPNGTFLAYAOQ 178
Db 167 QNNIYLKORPGDPFPQITFNGRENKIFNGIPDWYVEEMLPTKALWSPNGKFLAYAEF 226
Qy 179 NDETVPLIEVSFYDESLOPKTVRVPYKAGAVNPTVKFVNVNDSLSSVNTNATSIQIT 238
Db 227 NDKDIPVIAYSYSGDE--QYPRTNIPYKAGAKNVVRIPDIITTPAYVGPQ---EVP 281
Qy 239 APASMLIGDHYLCVDTWATERISLOWLRRIQNVSMVDCIDYDESSGRWNCVLVARQHTEM 298
Db 282 VPAMIASDYFSLWTWTERVCLQWLKRVQNVSVLSICDFREDWQWDCPKTQEHIEE 341
Qy 299 STTGWGRFRPSEPHFTLDGNSFYKIIISNEEGYRHCYFQIDKDKDCTFITKGTWEVIGIE 358
Db 342 SRTGWAGGFVSRPVFSYDAISYKIFSDXGKYHIHYKIDTVENAIQITSGKWEAINIF 401
Qy 359 ALTSYDLYISNEYKMGPGGRNLYKIQLSDY-TKVTCLSCELNPERCOYYSVSFSKBAKY 417
Db 402 RVTQDSLFSYSSNEPEEYFGRNRYIRISIGSYPPSKCVTCHLRKERCQYTYTASFSDYAKY 461
Qy 418 YQLRCSGPGPLVTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKLDLFIILNETKFWYQ 477
Db 462 YALVCYGGPISLTHDGRDQDEIKILEENKELENALKNTQLPKBEIKLEVDIEITLWYK 521
Qy 478 MILPPHDKSKYPLLLDVTYAGCSQKADTVFRLNWTATYLASTENIIVASFDGGRGSGYQG 537
Db 522 MILPPQFDRSKYPLLLQVYGGPSCSVRSVFAVNWLSYLSKSGVIALVDGRTAFQG 581
Qy 538 DKIMHAINRLGTFEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVVTSMVLSGSGVVF 597
Db 582 DKLLYAVYRKLGVYEVEDQITAVRKFTMGFIDEKRIAIWGSYGGYVVSLLASAGTGLF 641
Qy 598 KCGIAPVSRWEYSDVYTERYNGLTPEDNLDHYNSTVMSRAENFKOVEYLLIHGTA 657
Db 642 KCGIAPVSSWEYASVYTERFMGLPTKODNLEHYKNSTVMARAEYFRVNDVLLIHGTA 701
Qy 658 DDNVHFQSQASIKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKQCFSL 715
Db 702 DDNVHFQNSAQIAKALVNAQVDFQAMWYSQDNHGL-SGLSTNHLHYTHMTHFLKQCFSL 758

RESULT 4
US-11-151-601-20
; Sequence 20, Application US/11151601
; Publication No. US2006003413A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Libermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; TITLE OF INVENTION: AND PROTEASE FAMILY MEMBERS AND USES THEREOF
; FILE REFERENCE: MPI00-054PIRCPLOMIDIVM
; CURRENT APPLICATION NUMBER: US/11/151,601
; CURRENT FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
```

```
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20
; LENGTH: 882
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-11-151-601-20

Query Match 13.5%; Score 525; DB 7; Length 882;
Best Local Similarity 26.3%; Pred. No. 9e-34;
Matches 186; Conservative 108; Mismatches 277; Indels 136; Gaps 33;

Qy 99 PNNTQVWTVSPVGHKLAYVWNNDIYVKIEPLNPSYRITWTG-----KEDIYNGITDW 152
Db 206 PADPDWI-----AFIHSNDIWNISNIVTRBERRRUYVHNELANNEEDARSAGVATFV 256
Qy 153 YEEVFSAYSALWMSPNGTF-----LAYAQFNDTEVPLEIEYSFYSDESLOPKTVR 203
Db 257 LQEE-PRYSGYWMCPKAETTPSGGKTLRLIYENDESEVEIHH---VTSFMLETRADS 312
Qy 204 VPYKAGAVNPTVKFV--VNTDLSVTVNATSIQITAPASMLI-GDHYLCVDTWA---- 256
Db 313 FRYPKTGATNPKVTFKMSIEMIDAEGRIIDVDELQITQPPEILPEGVETIARAGWTPEGK 372
Qy 257 -----TOERISLQWL-----RRIQNY--SVMDICDYDESSGRW-N 288
Db 373 YAWSILLDRSQRLQVILISPELFIPEDDVMERQRLIESVPDSVPTLIIYEETDIWIN 432
Qy 289 C-----LVARQHIEMSTTGWGRFPESEPHFTLDGNSFYKIIIS--NEEGYRHI----- 334
Db 433 IHDIFHVFPQSHEEIEFIFASECKTGFRH-----LYKITSILKESKYRSGGGLPAP 485
Qy 335 CYPQIDKDOCTFTTKGTWEVIG-----IEALTSYDLYISNEYKMGPGGRNLYKI----- 384
Db 486 SDFKCPIKEBIAITSGEWEVLGRHGSNIQVDEVRLVYPEGT-KDSPLEHLYVVSYPNP 544
Qy 385 ----QLSD--YTKVTCLSCELNPERCOYYSVSFSKBAKYQLRCSGPGPLVTLHSSVND 438
Db 545 GEVTRUTDRGYSHSCCIS-----QHCDFFISKSNQKNPH---C-----VSLYKLSPEDD 592
Qy 439 KGLRVLED-NSALDKN--LQNVQMPSSKGLDFFILNETKF-WYQWILPPH-FDKSKYKPL 493
Db 593 PTCKTKEFWATILDSAGPLPDYTPPE---IFPESITGTGLYKMLYKPHDLQPKKYPTV 649
Qy 494 LDVIAGPCSQ-----KADTVFRLNWTATLASTENIIVASFDGGRGSGYQGDKIMHAINRR 547
Db 650 LFTYGGPQVLVNNRFGVKYFRLN---TLASLGYVWV-IDNRGSHRGLKFEAPKFKY 705
Qy 548 LGTFEVEDQIEAARQF-SKMGFVDNKRIAIWGSYGGYVTSVWLGSYGVPKCGIAPV 606
Db 706 MGQIEIDQVQVEGLQYLASRYDFIDLRVGHGWSYGGYLSLMAQMSQDIFRVAIAGPV 765
Qy 607 SRWEYSDVYTERYMGMLPTPEDNLDHYNSTVMSRAENF--KQVEYLLIHGTADDNVHFQ 664
Db 766 TLWIFYDTGTYERYMG--HPDQNEQGYLGSVAMQAEKFPSEPNRLLLHGHFLDENVHFA 823
Qy 665 QSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOIYTHMSHFIKO 711
Db 824 HTSILLSFLVRAGKPYDLQIYPOERHSIRVPESGEHYELHLLHYLQE 870
```



```
; PRIOR APPLICATION NUMBER: 60/227800
; PRIOR FILING DATE: 2000-08-25
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 1609
; SOFTWARE: CuraSeqlist version 0.1
; SEQ ID NO 1094
; LENGTH: 1436
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-453-372-1094

Query Match      2.9%; Score 112.5; DB 6; Length 1436;
Best Local Similarity 18.18; Pred. No. 0.83;
Matches 153; Conservative 109; Mismatches 273; Indels 309; Gaps 40;

Qy 55 YSISPDQFIL--LEYNVKQWRHSYASYDIYDLNKRQLITBERIPNNQWTVWSPVGH 112
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
19 YSLHPGEMVTDSTQPRIEKFCFENQVNLNFS-----KEKSLPGSNIDLQVSAASN 68
Qy 113 KLAYVWNNDIYVKLEPN-----LPSYRITWTCKED-----II 144
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
69 SLCALWAVDOSVLLRLNYGOLSAQTVYSQYLSRELHGYYFRGLNLEDGLKVPCLDEHIL 128
Qy 145 YNGITDWWYEEVFESAYSALWSPNGTFLAYAQFNDEVEPLIE-----YSFY 191
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
129 YNGI-----YYTP-----AWADFCKGYDLVDPQNNRIFQRQVTSFR 167
Qy 192 SDLSQVPTKVRVPYPKAGAVNPTVFFVNVNTDLSGVTVNATSIQ-----ITAPA 241
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
168 NITQLSQ---LISEPMFGDY-----WIVKRRNSRETHTQFAVRYVLPKFEVTNAPQ 219
Qy 242 SMLIG-DHYLCDYTWATQERISLOWLRRIQNSYVMDIC-DYDESSGRWNC-----LVARQ 294
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
220 TVTISDDEFQVDC-----AKYFQGPVQGETQIRVCREYFSSS-----NCEKNEICEQ 270
Qy 295 HIEWSTTGWGRFRSPHPTLQNSFY-----KIISNEGYRHRICVFQIDKKDCTPTIK 349
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
271 FIAQLENGCVSQIVNTKV-FQLYRSLGTFMTHVAVITESG-----TWQISEKTSVFIQ 325
Qy 350 --GTWEVIGIEALTSYLYIIS-----NEYKMGPGGRNLYKIQISD-----YT----- 390
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
326 LLGT---VNFENMDTFYRRGISYFGTLKFDSPDNNVPMVKNLQLELNDEFIGNYTTDENG 382
Qy 391 -----KVTCLSCLNPERC-----QYYSVSFSKEAKYY-----418
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
383 EAQFSIDTSIDFDPFNLKATYVRPESCYLPSWLTPOYLDAPH-LVSRFYSTRNSFLKIV 441
Qy 419 ----QLRCSGFLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSKKLDFIILNETKF 474
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
442 PEPKQLECNQO--KVTVVHYSLASE---AYEDDS-----NVKF 474
Qy 475 WYQMLPPHPDKKXKYPDLLD-----VYAGPCSQKADTVFRLNATWYLASTEN 522
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
475 FYLMWV-----KGAILLSGQKEIRKAWNGNFFPISISADLAPAAVLVYVTLHPSG 526
Qy 523 IIVASPDGRSGVGQDK-----IMAINRRLG 549
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
527 EIVAD-----SVRFQVDFCKFKHKNVIFKSNEQGLPGSNASLCLOAAPVLCALRAVDNRNL 582
Qy 550 TFEVEDQIEAARQFSKMGFDVNDKRIAIWGSYGYVTSMVLGSGS-----GVFKCGI 601
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
583 LLKSEQQLSAESVYNMVPSEI-----PYGVFYHGLNLDGKEDPCIPQRDMFVNLG 633
Qy 602 AVAPSRWEYDYSYTERYNGLPTPEDNLDHYNSTVMS-----RAENF 645
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
634 YVTPVSNYGGDIYNIVRNMLKV-FTNL-HYRKPEKIMQCVVFRLELHVASGIRGENA 691
Qy 646 KQVEYLLIHGTADD-----NVHFQSAQISKALVDGVDFQNMWYT---DEDHGI 692
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
692 DYVEQALIQVTRTNFPBTWMDLVSDSSGSANLSFLIPDTITIQWEASGFCVNGDVGFGI 751
Qy 693 ASST 696
||| : : :
```

```
Db 752 SSTT 755

RESULT 9
US-11-116-939-27
; Sequence 27, Application US/11116939
; Publication No. US20050265995A1
; GENERAL INFORMATION:
; APPLICANT: Stephen Tomlinson
; APPLICANT: Richard J. Quigg
; TITLE OF INVENTION: TISSUE TARGETED COMPLEMENT MODULATORS
; FILE REFERENCE: 19113.0115U2
; CURRENT APPLICATION NUMBER: US/11/116,939
; PRIOR FILING DATE: 2005-04-28
; PRIOR APPLICATION NUMBER: 60/565,907
; NUMBER OF SEQ ID NOS: 27
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 19
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence; note=synthetic
; OTHER INFORMATION: construct
US-11-116-939-27

Query Match      2.8%; Score 110; DB 7; Length 19;
Best Local Similarity 100.0%; Pred. No. 0.0035;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 695 STAQHIYTHMSHFIIKQCF 713
Db ||| : : : : : : : : : : : : : : : : : : : : : : : : : : : :
1 STAQHIYTHMSHFIIKQCF 19

RESULT 10
US-11-151-601-22
; Sequence 22, Application US/11151601
; Publication No. US20060003413A1
; GENERAL INFORMATION:
; APPLICANT: Milennium Pharmaceuticals, Inc.
; APPLICANT: Meyers, Rachel E.
; APPLICANT: Olandt, Peter J.
; APPLICANT: Kapeller-Liebermann, Rosana
; APPLICANT: Curtis, Rory A. J.
; APPLICANT: Williamson, Mark
; APPLICANT: Weich, Nadine
; TITLE OF INVENTION: NOVEL HUMAN PROTEIN KINASE, PHOSPHATASE,
; FILE REFERENCE: MP100-054PIRCPIOMNIDVIM
; CURRENT APPLICATION NUMBER: US/11/151,601
; PRIOR FILING DATE: 2005-06-13
; PRIOR APPLICATION NUMBER: US 10/170,789
; PRIOR FILING DATE: 2002-06-13
; PRIOR APPLICATION NUMBER: US 09/797,039
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: PCT/US01/06525
; PRIOR FILING DATE: 2001-02-28
; PRIOR APPLICATION NUMBER: US 60/186,061
; PRIOR FILING DATE: 2000-02-29
; PRIOR APPLICATION NUMBER: US 09/882,166
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: PCT/US01/19269
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: US 60/212,078
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 09/934,406
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26052
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/226,740
; PRIOR FILING DATE: 2000-08-21
```

; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 22
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-11-551-601-22

Query Match 2.8%; Score 110; DB 7; Length 72;
Best Local Similarity 48.3%; Pred. No. 0.022;
Matches 29; Conservative 3; Mismatches 28; Indels 0; Gaps 0;

QY 524 IVASFDCGRSGYQCKIMHAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIATWGSYGG 583
DB 11 IYAVVDIRGGEGYQKWEAGTRKLKKNFDFIAAAAYLSKLYTSPKRIATPGSNGG 70

RESULT 11

US-11-077-550-157
; Sequence 157, Application US/11077550
; Publication No. US2005024435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22
; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 157
; LENGTH: 877
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-157

Query Match 2.8%; Score 109.5; DB 7; Length 877;
Best Local Similarity 19.5%; Pred. No. 0.73;
Matches 144; Conservative 91; Mismatches 259; Indels 243; Gaps 32;

QY 30 NAEYGNSSV---FLENSTDFEGHISINDYSISPDGQFILLEYNVVKQWRHSYASYDIYD 86
DB 250 NAYEYSGLEVSFEELRTFG--GHDAKFIDSLQENEFRLYYNKKF-----DIAS 297

QY 87 -LNKRQLITEERIPNNQTQWTVSPGVGHKLAYVWNNDIYVKIBPNLPSYRITWTGKEDIY 145
DB 298 TLNKAQSIV-----GTTASIQYKQNVFKYKYLLEDTSIGKFSVDKLF-----DKLY 344

QY 146 NGITDWYEEVSAYSALWSPNGFLAY--RQFNDTEVPLLEYSPYDESLOYKPTVR 203
DB 345 KMLTE-IYEDNFVKFKV--LNRKTYLNFDPKAVFINIVPKVNYTIYDGNLRLNTL-- 399

QY 204 VPYPKAGAVNPTVKFFVNTDSLSSVTNATSIQITAPASMLIGHY--LCDVTWATQERI 261
DB 400 -----AANFNGQNTENNWNFTKLNFTG-----LPEFYKLLCVRGIIITKTK 442
QY 262 SL-QMLRRIQYNSVMDICDYDESSGRWNCCLVARQHIEMSTTGWVGRPRPSPHPHTLQNS 320
DB 443 SLDKGYNKIRGRALNDLC-----IKVNNMDLFPSPSEDNFTNDLNK 483
QY 321 FYKIISN---EEGVRHICYFQIDKDOCTEITKTGWEVIGIEALTSVYL--YYISNEYKGM 375
DB 484 GBEITSDTNEAAEENISLDLIQOYILTFNFDNEPENISIEENSSDIIGOLELMPNTERF 543
QY 376 PGRNLYKIQISDYTKVTCISCEINPERCOYVSFSKEAKYQLRCSGPGCLPLYTLHSS 435
DB 544 PNGK---KYELDXYTWFHYL-----RAQEPHEKSRIA-----LTNS 577
QY 436 VNDKGRVLVEDNSALDKMLQNVOMPSKKLDPIILNETKFWYQWILPDPHFKSKYPLLLD 495
DB 578 VNEALLN-----PSRVYTFE----- 592
QY 496 VYAGPCSQKADTVPRLNWAT-----YLASTENIIVASPDGRSGVQGDKIMH----- 542
DB 593 -----SSDYVKKVKNKATEAAMFLGWVEQLVYDFTDETSEVSTTDKIADITIIIPYIG 644
QY 543 -AINRRLGTFEVEDQIEAARQFS-----KMGFVNDKRIATWGSYGYVTSMVLGSGSV 596
DB 645 PALN--IGNMLYKDDFVGALIFSGAVILLFEIPIALPVLGTALVSYIANKVLTVOT-- 700
QY 597 FKCGIAPVPSRWYYSVYTERYMGLEPTEDNLHYRNSVTMSRAENFKQVEYLLHGT 656
DB 701 ----IDNALSKRNEKWDVY-----KYIVTNWL 724
QY 657 ADDNVH-----FQSAQISKALVDVGVDFQAMWYTDHGG-----IASSTAH 698
DB 725 AKYNTQIDILIRKQKKEALENQAETKAI-----INYQYNTYEBEKNINFNIDDLSSKLN 780
QY 699 QHI---YTHMSHFQK 712
DB 781 ESINKAWINKELNQC 797

RESULT 12

US-10-647-956A-2
; Sequence 2, Application US/10647956A
; Publication No. US2005025187A1
; GENERAL INFORMATION:
; APPLICANT: French-Constant, Richard
; APPLICANT: Bowen, David
; APPLICANT: Rocheleau, Thomas
; APPLICANT: Waterfield, Nicholas
; TITLE OF INVENTION: DNA SEQUENCES FROM PHOTORHABDUS LUMINESCENS
; FILE REFERENCE: 61645
; CURRENT APPLICATION NUMBER: US/10/647,956A
; CURRENT FILING DATE: 2003-08-26
; PRIOR APPLICATION NUMBER: CURRENT APPLICATION NUMBER: US/09/817,514
; PRIOR FILING DATE: CURRENT FILING DATE: 2000-03-26
; PRIOR APPLICATION NUMBER: US 60/191806
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 2516
; TYPE: PRT
; ORGANISM: Photorhabdus luminescens
US-10-647-956A-2

Query Match 2.8%; Score 109; DB 6; Length 2516;
Best Local Similarity 18.5%; Pred. No. 3.4;
Matches 137; Conservative 89; Mismatches 205; Indels 310; Gaps 39;

QY 20 YKQENNILV--FNAE-----YGNSS-----VFLNSTFDFEGHISINDYSISPDGQFILLE 67
DB 1239 YQGEDTLLVMFYNQDTLDSYKNASMQGLYIFADMASKDMTPQSNVYR---DNSYQQFD 1295

```

QY 68 YNVYKQWRHSYASYDI-YDLNKRQ-----LITEERIPNNTQWVTWSPVGHKLAY 116
Db 1296 TNNVRNNRYAEDEYIEPSSVSRKQYGMWDYLSMVGNDIPT-----INYKAA- 1345
QY 117 VMNDIVVKEPNLPSYRITWTGKEDIYNGITDWVE-----EEVFSA 160
Db 1346 --SDLKIIYSPKL-----RIIHG-----YEGQRNQCNMNKKYKGLGDKFIV 1387
QY 161 YSALWSPNGT---FLAYAQFNDTEVPLIEYS-----FYSDLSLOYPKTVR-- 203
Db 1388 YTSLVGNPNNSNKLMEY-----PVYQYSGNTSGLNQRLLFHRDITI--YPSKVEAW 1437
QY 204 VPPKAGAVNPT-----VKFFVNTDSLSVNTNATS-IQIT---APAS 242
Db 1438 IPGAKRSLTNQNAAGDDYATDSLNLKDDLLQYIFMTDSKGTATDVSGPVEINTAISPAK 1497
QY 243 MLI-----GDHYLCDVTWATERISLQ-----WLRRIQNVSMDI- 277
Db 1498 VQIIVKAGGK---EQFTADKDVSIQPSFDEMNQYQFNALEIDGSGLNFINNASIDVT 1554
QY 278 -----CDYDESS-----GRMNCLVARQHI 296
Db 1555 FTAPAEGRKLGYESFSIPVTLKVSTDNALTLLHNENGAQYMQWQSVYRTRLNLFARQLV 1614
QY 297 EMSTTG-----WVGRFRPSPHFTLDGNSFY-----KLIISN 327
Db 1615 ARATTTGDTILSMETQIQEPQL---GKGYATFVIPPYNLSTHGDERNFKLYIKHVDN 1671
QY 328 EGVYRHCY-----FOID-----KKDCTF---ITKGTW-----E 353
Db 1672 NS---HIYSGQLTDTNITLFIPLDDVPLNQDYHAKVYMTFKKSPSDGTWNGPHFVRD 1728
QY 354 VIGIEAL-----TSDYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPER 403
Db 1729 DKGIVTNPKSILTHPESVNVLANISSEPMDFSGANSLYFWELFYITPMLVAQRLLEHQ- 1787
QY 404 QYYSVSFSKEAYQYLCRSGGGLPLYTLHSSVNDKGLRV---LEDNS----- 448
Db 1788 -----NFDENRNLKYVWSPSG---YIVHGQIQNYQWVRLPLEDTSWNSDPLDSVDPD 1838
QY 449 ---ALDKMLQNVQMPKSLDFI1-----LNETKFWYQWILPPHFDKSKYK 491
Db 1839 AVAQHPMHYKVSTFWRTLDLLIARGDHAYRQLERDTLNEAKWYQWAL--HLLGDKPYL 1896
QY 492 LLLDVYAGPCSQKADTVFRLN 512
Db 1897 PLSTTWSDPRLDRAADITQN 1917

```

RESULT 13

```

US-11-077-550-32
; Sequence 32, Application US/11077550
; Publication No. US20050244435A1
; GENERAL INFORMATION:
; APPLICANT: Shone, Clifford Charles
; APPLICANT: Quinn, Conrad Padraig
; APPLICANT: Foster, Keith Alan
; APPLICANT: Chaddock, John
; APPLICANT: Marks, Philip
; APPLICANT: Sutton, J. Mark
; APPLICANT: Stancombe, Patrick
; APPLICANT: Wayne, Jonathan
; TITLE OF INVENTION: Recombinant Toxin Fragments
; FILE REFERENCE: 1581.0130004
; CURRENT APPLICATION NUMBER: US/11/077,550
; CURRENT FILING DATE: 2005-03-11
; PRIOR APPLICATION NUMBER: 10/241,596
; PRIOR FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: 09/255,829
; PRIOR FILING DATE: 1999-02-23
; PRIOR APPLICATION NUMBER: PCT/GB97/02273
; PRIOR FILING DATE: 1997-08-22

```

```

; PRIOR APPLICATION NUMBER: 08/782,893
; PRIOR FILING DATE: 1996-12-27
; PRIOR APPLICATION NUMBER: GB9625996.5
; PRIOR FILING DATE: 1996-12-13
; PRIOR APPLICATION NUMBER: GB9617671.4
; PRIOR FILING DATE: 1996-08-23
; NUMBER OF SEQ ID NOS: 179
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 32
; LENGTH: 866
; TYPE: PRT
; ORGANISM: Clostridium botulinum
US-11-077-550-32

Query Match 2.8%; Score 107; DB 7; Length 866;
Best Local Similarity 19.6%; Pred. No. 1.1;
Matches 144; Conservative 89; Mismatches 261; Indels 242; Gaps 31;

QY 30 NAEYGNSSV---PLENSTFDERGHSINDYSISPDGOFILLEYNVYKQWRHSYASYDIYD 86
Db 248 NAYYEMSGLEVSEELRTFG--GHDAKFIDSQENEFRLYYNKFK-----DIAS 295
QY 87 -LNKROLITEERIPNNTQWVTWSPVGHKLAYVMNNDIYVKIEPNLPSYRITWTGKEDIY 145
Db 296 TLNKAISV-----GTASLQYMKNVFKEKILLSEDTSGKFSVDLKF-----DKLY 342
QY 146 NGITDWMVEEVFSAYSALWSPNGTFLAY--AQFNDTEVPLIEYSFYSDLSLOYPKTVR 203
Db 343 KMLTE-IYTEDNFVKPFKV--LNRKTYLNFDAVKINIVPKVNYTYIDGFNLRLNTNL-- 397
QY 204 VYPKAGAVNPTVKFPVNTDSLSVNTNATSQITAPASMLIGDHY--LCDVTWATERI 261
Db 398 -----AANFNGQNTENNMTKLKNTG-----LFEFYKLLCVRGIITSKTK 440
QY 262 SLQWLRRIQNVSMIDICDYDESSGRMNCLVARQHIEMSTTGWGRFRPSPHFTLDGNSF 321
Db 441 SLDDDDKGYNKALNDLC-----IKVNMWDLFPSDEDNFTDLNKG 481
QY 322 YKIISN---EEGVRHCYFOIDKKOCTFITKGTWIEGIEALTSYDL--YYISNEYKMG 376
Db 482 EBITSDTNEAAEENISLDLIQQYVLTNFNDEPENISIENTLSSDIIGOLELMPNIERFP 541
QY 377 GERNLYKIQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQYLCRSGGGLPLYTLHSSV 436
Db 542 NGK---KYELDKYTMFHYL-----RAQEFEGKSRIA-----LTNSV 575
QY 437 NDKGLRVLEDNSALDKMLQNVQMPKSLDFI1LNETKFWYQWILPPHFDKSKYKPLLDV 496
Db 576 NEALLN-----PSRVYTFP----- 589
QY 497 YAGPCSQKADTVFRLNWTAT-----YLASTENIIVASPDGRSGVYQGGKIMH----- 542
Db 590 -----SSDYVKVKNKATEAAMFLGWVEQLVYDFDTESEVSTTDKIADITIIIPYIGP 642
QY 543 AJNRRLTGFEVEDQIEAARQFS-----KMGFVDNKRITAWG--WSYGGYVTSWVLGSGGVF 597
Db 643 ALN--IGNMLYKDDFVGALIFSGAVILLEFIPETALPVLGTFAVLSYIANKVLTQV--- 697
QY 598 KCGIAPVPSVRWEYDVSVYTERYMGLTPTPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTA 657
Db 698 ---IDNALSKEKDEWVY-----KIVITNWL 722
QY 658 DDNVH-----FQSAQISKALVDVGVDVQAMWYTDHGH-----IASSTAHO 699
Db 723 KVNTQIDILIRKKMKEALENQAETRAI----INYQVNYTEBEKKNINFINIDLSKLNE 778
QY 700 HI----YTHMSHFQK 712
Db 779 SINKAMININKFLNQC 794

RESULT 14
US-11-077-550-155

```

Sequence 155, Application US/11077550
Publication No. US20050244435A1
GENERAL INFORMATION:
APPLICANT: Shone, Clifford Charles
APPLICANT: Quinn, Conrad Padraig
APPLICANT: Foster, Keith Alan
APPLICANT: Chaddock, John
APPLICANT: Marke, Philip
APPLICANT: Sutton, J. Mark
APPLICANT: Stancombe, Patrick
APPLICANT: Wayne, Jonathan
TITLE OF INVENTION: Recombinant Toxin Fragments
FILE REFERENCE: 1581.0130004
CURRENT APPLICATION NUMBER: US/11/077,550
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: 10/241,596
PRIOR FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: 09/255,829
PRIOR FILING DATE: 1999-02-23
PRIOR APPLICATION NUMBER: PCT/GB97/02273
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 08/782,893
PRIOR FILING DATE: 1996-12-27
PRIOR APPLICATION NUMBER: GB9625996.5
PRIOR FILING DATE: 1996-12-13
PRIOR APPLICATION NUMBER: GB9617671.4
PRIOR FILING DATE: 1996-08-23
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn version 3.1
SEQ ID NO 155
LENGTH: 871
TYPE: PRT
ORGANISM: Clostridium botulinum
US-11-077-550-155

Query Match 2.8%; Score 107; DB 7; Length 871;
Best Local Similarity 19.3%; Pred. No. 1.1;
Matches 147; Conservative 94; Mismatches 263; Indels 258; Gaps 34;

16 HEVLYKQE-----NNILVFNAEYGNSSVPLENS-----TFDEF-----GHSINDYSISPD 60
223 HELIYAGHRLYGIAINPNRVKIV---NTNAYEMSGLEVSFEELRTFGCHDAKFDISLOE 279
61 GQFILLEVYVQWRHSYTSYDIYD-LNKRQLITEERIPNNTQVWTSVPVGHKLAYVNN 119
280 NEFRLYYYNKKF-----DIASLTKAKSIV-----GTTASLQYMKVNFKEKYL 324
120 NDIYVYKIEPNLSYRITWTGKEDIYNGITDWTYEEVFSAYSALWSPNGTFLAY--AQ 177
325 EDTSGKFSVDKUKF-----DKLYKMLTE-IYTFEDNFVFFPKV--LNRKTYLNFDPKAV 373
178 FNDTEVPLIEYGFYSDLSQYVKTVRVPYKAGAVNPTVKFFVNVNTDLSLSSVTNATSIQI 237
374 FKINIVPKVNYIYOGFNLRTNLT-----AANFNGQTEINNNWFTKLNKFTG--- 421
238 TAPASMLIGDHY--LCDVTWATQERISLQWLRRIQYNSVMDICDYDESSGRWNCVLARQH 295
422 -----LFEFYKLLCVRGIIITSKTSLD---KGYNKALNDLC----- 454
296 IEMSTTGWGRPRPSPHPTLDGNSFYKIISN---BEGYRHICYFQDKDCTFITKGTW 352
455 --IKVNNMDLFFSPSEDNFTDLNKGEBITSDTIEAAEENISLDLIQQYYLTENFQNEP 512
353 EVIGIEALTSYDL--YYISNEYKMGPGGRNLYKQLSDYTKVTCGLSCELNPERCOYYSVS 410
513 ENISIEJSSDIIQGLELWPNIERPNKG---KYELDKYTFMFWHL-----RAQBEFHG 562
411 FSKEAKYIQLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFTILN 470
563 KSRIA-----LTVSNVEALLN-----PSRVVYTF-- 586
471 ETKFYQMLPPLPHFDKSKYPLLLLDVYAGPCSQKADTVFRLNWTAT-----YLASTENTIVA 526

US-11-077-550-163

Sequence 163, Application US/11077550
Publication No. US20050244435A1
GENERAL INFORMATION:
APPLICANT: Shone, Clifford Charles
APPLICANT: Quinn, Conrad Padraig
APPLICANT: Foster, Keith Alan
APPLICANT: Chaddock, John
APPLICANT: Marke, Philip
APPLICANT: Sutton, J. Mark
APPLICANT: Stancombe, Patrick
APPLICANT: Wayne, Jonathan
TITLE OF INVENTION: Recombinant Toxin Fragments
FILE REFERENCE: 1581.0130004
CURRENT APPLICATION NUMBER: US/11/077,550
CURRENT FILING DATE: 2005-03-11
PRIOR APPLICATION NUMBER: 10/241,596
PRIOR FILING DATE: 2002-09-12
PRIOR APPLICATION NUMBER: 09/255,829
PRIOR FILING DATE: 1999-02-23
PRIOR APPLICATION NUMBER: PCT/GB97/02273
PRIOR FILING DATE: 1997-08-22
PRIOR APPLICATION NUMBER: 08/782,893
PRIOR FILING DATE: 1996-12-27
PRIOR APPLICATION NUMBER: GB9625996.5
PRIOR FILING DATE: 1996-12-13
PRIOR APPLICATION NUMBER: GB9617671.4
PRIOR FILING DATE: 1996-08-23
NUMBER OF SEQ ID NOS: 179
SOFTWARE: PatentIn version 3.1
SEQ ID NO 163
LENGTH: 873
TYPE: PRT
ORGANISM: Clostridium botulinum
US-11-077-550-163

Query Match 2.8%; Score 107; DB 7; Length 873;
Best Local Similarity 19.3%; Pred. No. 1.1;
Matches 147; Conservative 94; Mismatches 263; Indels 258; Gaps 34

16 HEVLYKQE-----NNILVFNAEYGNSSVPLENS-----TFDEF-----GHSINDYSISPD 60
225 HELIYAGHRLYGIAINPNRVKIV---NTNAYEMSGLEVSFEELRTFGCHDAKFDISLOE 281
61 GQFILLEVYVQWRHSYTSYDIYD-LNKRQLITEERIPNNTQVWTSVPVGHKLAYVNN 119
282 NEFRLYYYNKKF-----DIASLTKAKSIV-----GTTASLQYMKVNFKEKYL 326
120 NDIYVYKIEPNLSYRITWTGKEDIYNGITDWTYEEVFSAYSALWSPNGTFLAY--AQ 177
327 EDTSGKFSVDKUKF-----DKLYKMLTE-IYTFEDNFVFFPKV--LNRKTYLNFDPKAV 375
178 FNDTEVPLIEYGFYSDLSQYVKTVRVPYKAGAVNPTVKFFVNVNTDLSLSSVTNATSIQI 237

[illegible]

Search completed: January 28, 2006, 00:05:16
Job time : 15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2006 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 30, 2006, 14:06:40 ; Search time 163 Seconds
(without alignments)

3099.130 Million cell updates/sec

Title: US-10-659-055-1_COPY_51_766

Perfect score: 3877

Sequence: 1 NTVRLKLYSLRWISWDHELYL.....AQHIYTHMSHFIRKQCFSLP 716

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

UniProt_05.80.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3877	100.0	766	1 DPP4_HUMAN	P27487 h dipeptidyl
2	3877	100.0	766	2 Q53YN1_HUMAN	Q53tn1 homo sapien
3	3848	99.3	765	2 Q5R7G7_PONPY	Q5r7g7 pongo pygma
4	3531	91.1	765	1 DPP4_BOVIN	P81425 b dipeptidyl
5	3501	90.3	765	1 DPP4_FELCA	Q9n217 f dipeptidyl
6	3489	90.0	766	1 DPP4_PIG	P22411 s dipeptidyl
7	3354.5	86.5	767	1 DPP4_RAT	P14740 r dipeptidyl
8	3340	86.1	760	1 DPP4_MOUSE	P28843 m dipeptidyl
9	2490.5	64.2	751	2 Q75883_AGKHB	Q75883 agkistrodon
10	2489.5	64.2	751	2 Q75882_AGKHB	Q75882 agkistrodon
11	2478	63.9	759	2 Q52181_CHICK	Q52181 gallus gall
12	2309	59.6	737	2 Q641D6_XENLA	Q641d6 xenopus lae
13	2292.5	59.1	748	2 P70092_XENLA	P70092 xenopus lae
14	2145	55.3	760	2 Q53TP5_HUMAN	Q53tp5 homo sapien
15	2141.5	55.2	761	2 Q8R492_RAT	Q8r492 rattus norv
16	2138	55.1	760	1 SEPR_HUMAN	Q12884 homo sapien
17	2136.5	55.1	761	1 SEPR_MOUSE	P97331 mus musculus
18	2103	54.2	755	2 Q91651_XENLA	Q91651 xenopus lae
19	1966.5	50.7	860	2 Q4RQJ4_TETNG	Q4rqj4 tetraodon n
20	1819.5	46.9	704	2 Q4S309_TETNG	Q4s309 tetraodon n
21	1216	31.4	796	1 DPP10_HUMAN	Q8n608 homo sapien
22	1207	31.1	796	1 DPP10_RAT	Q8q629 rattus norv
23	1192.5	30.8	796	2 Q6INB7_XENLA	Q6inb7 xenopus lae
24	1182	30.5	797	1 DPP10_MOUSE	Q6nxx7 mus musculus
25	1117	28.8	865	1 DPP6_HUMAN	P42658 homo sapien
26	1111.5	28.7	731	2 Q7PSE9_ANOGA	Q7pse9 anopheles g
27	1109	28.6	803	1 DPP6_PANTR	Q5i650 pan troglod
28	1106	28.5	803	2 Q9QW8_9MURI	Q9qcw8 rattus sp.
29	1104	28.5	859	1 DPP6_RAT	P46101 rattus norv
30	1101	28.4	745	2 Q9VNM2_DROME	Q9vnm2 drosophila
31	1094	28.2	863	1 DPP6_BOVIN	P42659 bos taurus

32	1077	27.8	801	2 Q5U4C2_MOUSE	Q5u4c2 mus musculus
33	1077	27.8	803	2 Q80VM5_MOUSE	Q80vm5 mus musculus
34	1077	27.8	804	1 DPP6_MOUSE	Q92218 mus musculus
35	1057.5	27.3	815	2 Q4RK63_TETNG	Q4rk63 tetraodon n
36	1052	27.1	724	2 Q51895_9PRIM	Q51895 saimiri bol
37	1051	27.1	604	2 Q538S4_HUMAN	Q538s4 homo sapien
38	1047.5	27.0	439	2 Q6PG25_XENLA	Q6pg25 xenopus lae
39	1045	27.0	801	2 Q4RK66_TETNG	Q4rk66 tetraodon n
40	1030	26.6	802	2 Q9VUH1_DROME	Q9vuh1 drosophila
41	1025.5	26.5	935	2 Q9VMB4_DROME	Q9vmb4 drosophila
42	998	25.7	775	2 Q5J6J3_TIRIRU	Q5j6j3 trichophyto
43	996	25.7	842	2 Q4RVF8_TETNG	Q4rvf8 tetraodon n
44	975	25.1	874	2 Q7SI80_EMENI	Q7si80 emericeella
45	967	24.9	880	2 Q5B934_EMENI	Q5b934 aspergillus

ALIGNMENTS

RESULT 1
DPP4_HUMAN
ID DPP4_HUMAN STANDARD; PRT; 766 AA.
AC P27487;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 13-SEP-2005 (Rel. 48, Last annotation update)
DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
IV) (T-cell activation antigen CD26) (TP103) (Adenosine deaminase
DE complexing protein 2) (ADABP) [Contains: Dipeptidyl peptidase 4
DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
GN Name=DPP4; Synonyms=ADCP2, CD26;
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Theria; Eumarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Liver;
RX MEDLINE=92329551; PubMed=1352704; DOI=10.1016/0167-4781(92)90036-Y;
RA Misumi Y., Hayashi Y., Arakawa F., Ikehara Y.;
RT "Molecular cloning and sequence analysis of human dipeptidyl peptidase
RL IV, a serine proteinase on the cell surface.";
RL Biochim. Biophys. Acta 1131:333-336(1992).
[2]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Placenta;
RX MEDLINE=95012454; PubMed=7927537;
RA Abbott C.A., Baker E., Sutherland G.R., McCaughan G.W.;
RT "Genomic organization, exact localization, and tissue expression of
RL the human CD26 (dipeptidyl peptidase IV) gene.";
RL Immunogenetics 40:331-338(1994).
[3]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=peripheral blood;
RX MEDLINE=92325476; PubMed=1352530;
RA Tanaka T., Camerini D., Seed B., Torimoto Y., Dang N.H., Kameoka J.,
RA Dahlberg H.N., Schlossman S.F., Morimoto C.;
RT "Cloning and functional expression of the T cell activation antigen
RL CD26.";
RL J. Immunol. 149:481-486(1992).
[4]
RP ERRATUM.
RX MEDLINE=93171637; PubMed=8094732;
RA Tanaka T.;
RL J. Immunol. 150:2090-2090(1993).
[5]
RP NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=prostate, and Uterus;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,

RA Alteschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carinci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Craythorn J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[6]
RN NUCLEOTIDE SEQUENCE OF 1-551.
RP TISSUE=Colon;
RX MEDLINE=92165847; PubMed=1347043;
RA Darmoul D., Lacasa M., Baricault L., Marguet D., Sapin C., Trotot P.,
RA Barbat A.;
RT "Dipeptidyl peptidase IV (CD 26) gene expression in enterocyte-like
RT colon cancer cell lines HT-29 and Caco-2. Cloning of the complete
RT human coding sequence and changes of dipeptidyl peptidase IV mRNA
RT levels during cell differentiation.";
RL J. Biol. Chem. 267:4824-4833(1992).
[7]
RN NUCLEOTIDE SEQUENCE OF 545-766.
RP TISSUE=Colon;
RX MEDLINE=91024044; PubMed=1977364;
RA Darmoul D., Lacasa M., Chantret I., Swallow D., Trugnan G.;
RT "Isolation of a cDNA probe for the human intestinal
RT dipeptidylpeptidase IV and assignment of the gene locus DPP4 to
RT chromosome 2.";
RL Ann. Hum. Genet. 54:191-197(1990).
[8]
RN NUCLEOTIDE SEQUENCE OF 1-31.
RX MEDLINE=96067599; PubMed=7487939;
RA Boehm S.K., Gum J.R. Jr., Erickson R.H., Hicks J.W., Kim Y.S.;
RT "Human dipeptidyl peptidase IV gene promoter: tissue-specific
RT regulation from a TATA-less GC-rich sequence characteristic of a
RT housekeeping gene promoter.";
RL Biochem. J. 311:835-843(1995).
[9]
RN PROTEIN SEQUENCE OF 1-22, AND TISSUE SPECIFICITY.
RP PubMed=1677636;
RA Gorvel J.P., Ferrero A., Chambraud L., Rigal A., Bonicel J.,
RA Maroux S.;
RT "Expression of sucrase-isomaltase and dipeptidylpeptidase IV in human
RT small intestine and colon.";
RL Gastroenterology 101:618-625(1991).
[10]
RN PARTIAL PROTEIN SEQUENCE.
RP TISSUE=Kidney;
RX MEDLINE=93210468; PubMed=8096237; DOI=10.1084/jem.177.4.1135;
RA Morrison M.E., Vijayasekaradi S., Engelstein D., Albino A.P.,
RA Houghton A.N.;
RT "A marker for neoplastic progression of human melanocytes is a cell
RT surface ectopeptidase.";
RL J. Exp. Med. 177:1135-1143(1993).
[11]
RN X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 38-766 IN COMPLEX WITH
RP INHIBITOR, AND HOMODIMERIZATION.
RX PubMed=12832764; DOI=10.1107/S0907444903010059;
RA Oefner C., D'Arcy A., Mac Sweeney A., Pierau S., Gardiner R.,
RA Dale G.E.;
RT "High-resolution structure of human apo dipeptidyl peptidase IV/CD26
RT and its complex with 1-[(2-[(5-iodopyridin-2-yl)amino]-ethyl)amino]-
RT acetyl]-2-cyano-(S)-pyrrolidine.";
RL Acta Crystallogr. D 59:1206-1212(2003).
[12]
RN X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 35-771, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12646248; DOI=10.1016/S0006-291X(03)00258-4;
RA Hiramatsu H., Inoko K., Higashiyama Y., Fukushima C., Shima H.,
RA Sugiyama S., Yano K., Yamamoto A., Shimizu R.;
RT "The structure and function of human dipeptidyl peptidase IV,
RT possessing a unique eight-bladed beta-propeller fold.";
RL Biochem. Biophys. Res. Commun. 302:849-854(2003).
[13]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12483204; DOI=10.1038/nsb882;
RA Rasmussen H.B., Branner S., Wiberg F.C., Wagtmann N.;
RT "Crystal structure of human dipeptidyl peptidase IV/CD26 in complex
RT with a substrate analog.";
RL Nat. Struct. Biol. 10:19-25(2003).
[14]
RN X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS) OF 39-766, HOMODIMERIZATION, AND
RP N-GLYCOSYLATION SITES.
RX PubMed=12906826; DOI=10.1016/S0969-2126(03)00160-6;
RA Thoma R., Loeffler B., Stihle M., Huber W., Ruf A., Hennig M.;
RT "Structural basis of proline-specific exopeptidase activity as
RT observed in human dipeptidyl peptidase-IV.";
RL Structure 11:947-959(2003).
CC -!- FUNCTION: Removes N-terminal dipeptides sequentially from the
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline. Plays a role in T cell activation.
CC -!- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -!- SUBUNIT: Homodimer or heterodimer with Seprase (PAP).
CC -!- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -!- TISSUE SPECIFICITY: Expressed in the poorly differentiated crypt
CC cells of the small intestine as well as in the mature villous
CC cells. Expressed at very low levels in the colon.
CC -!- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -!- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.

CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.

DR EMBL; U13735; AAB60646.1; Genomic DNA.
DR EMBL; U13710; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13711; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13712; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13713; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13714; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13715; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13716; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13717; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13718; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13719; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13720; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13721; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13722; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13723; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13724; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13725; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13726; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13727; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13728; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13729; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13730; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13731; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13732; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13733; AAB60646.1; JOINED; Genomic DNA.
DR EMBL; U13734; AAB60646.1; JOINED; Genomic DNA.

DR EMBL; M74777; AA51943.1; -; mRNA.
DR EMBL; BC013329; AAH13329.2; -; mRNA.

Query Match 100.0%; Score 3877; DB 1; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.6e-256;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHEHYLYKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
DB 51 NTYRLKLYSLRWISDHEHYLYKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
QY 61 GQFILLEVYVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWNN 120
DB 111 GQFILLEVYVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
DB 231 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
QY 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIEWST 300
DB 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIEWST 350
QY 301 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVEIGIEAL 360
DB 351 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVEIGIEAL 410
QY 361 TSDYLYIISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
DB 411 TSDYLYIISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
QY 421 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDPIILNETKFWYQWIL 480
DB 471 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDPIILNETKFWYQWIL 530
QY 481 PPHFDKSKKYPILLDVTAGPCSKADTVFRLNWTATYLASTENIIVASFDFGRSGYQGDKI 540
DB 531 PPHFDKSKKYPILLDVTAGPCSKADTVFRLNWTATYLASTENIIVASFDFGRSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNKRIAIWGSYGGYVTSWMLGSGGVFKCG 600
DB 591 MHAINRRLGTFEVEDQIEAARQFSKMGFVNKRIAIWGSYGGYVTSWMLGSGGVFKCG 650
QY 601 IAVAPSRWEYSDVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 660
DB 651 IAVAPSRWEYSDVYTERYMGVLPEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN 710
QY 661 VHFQSSAQISKALVDVGVDFQAMWYTDHGHIASTAHQIYTHMSHFIKQCFSLP 716
DB 711 VHFQSSAQISKALVDVGVDFQAMWYTDHGHIASTAHQIYTHMSHFIKQCFSLP 766

RESULT 2

Q53TN1 HUMAN
ID Q53TN1_HUMAN PRELIMINARY; PRT; 766 AA.
AC Q53TN1;
DT 13-SEP-2005 (TrEMBLrel. 31, Created)
DT 13-SEP-2005 (TrEMBLrel. 31, Last sequence update)
DT 13-SEP-2005 (TrEMBLrel. 31, Last annotation update)
DE Hypothetical protein DPP4.
GN Name=DPP4;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Ozanich A., Stoneking T., Hawkins M., Sapetti L.,

RT "The sequence of Homo sapiens BAC clone RP11-178A14.";
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RA NUCLEOTIDE SEQUENCE.
RP Waterston R.H.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RA NUCLEOTIDE SEQUENCE.
RP Waterston R.;
RL Submitted (SEP-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RA NUCLEOTIDE SEQUENCE.
RP Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
RN EMBL; AC008063; AA93179.1; -; Genomic_DNA.
DR SMR; O53TN1; 39-766.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0003824; F:catalytic activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR PROSITE; PS00708; PRO_ENOPEP_SER; 1.
DR Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 766 AA; 88279 MW; 5FBA2C6662D6117 CRC64;

Query Match 100.0%; Score 3877; DB 2; Length 766;
Best Local Similarity 100.0%; Pred. No. 3.6e-256;
Matches 716; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 NTYRLKLYSLRWISDHEHYLYKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 60
DB 51 NTYRLKLYSLRWISDHEHYLYKQENNLVFNNAEYGNSSVFLNSTFDFGHSINDYSISPD 110
QY 61 GQFILLEVYVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWNN 120
DB 111 GQFILLEVYVYKQWRHSYTSYDIYDLNKRQLITEERIPNNTQVWTSVPVGHKLAYVWNN 170
QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 180
DB 171 DIYVKIEPNLPSYRITWTGKEDIYNGITDWWYEEVFSAYSALWSPNGTFLAYAFND 230
QY 181 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 240
DB 231 TEVPLIEYSFYSDSLQYPKTVRVPYKAGAVNPTVKFFVNTDLSLSSVTNATSIQITAP 290
QY 241 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIEWST 300
DB 291 ASMLIGDHYLCVDTWATQERISLQWLRRIQNYSDICDYDESSGRWNCVLAHQHIEWST 350
QY 301 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVEIGIEAL 360
DB 351 TGWVGRFRPSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTFITKGTWVEIGIEAL 410
QY 361 TSDYLYIISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 420
DB 411 TSDYLYIISNEYKMGPGGRNLYKQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL 470
QY 421 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDPIILNETKFWYQWIL 480
DB 471 RCGSPGLPLTYLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSLDPIILNETKFWYQWIL 530
QY 481 PPHFDKSKKYPILLDVTAGPCSKADTVFRLNWTATYLASTENIIVASFDFGRSGYQGDKI 540
DB 531 PPHFDKSKKYPILLDVTAGPCSKADTVFRLNWTATYLASTENIIVASFDFGRSGYQGDKI 590
QY 541 MHAINRRLGTFEVEDQIEAARQFSKMGFVNKRIAIWGSYGGYVTSWMLGSGGVFKCG 600

Db	591	MHAINRRLGTFFVEVDQIEAARQFSKMGFVDNKRILAIWGSYGGVYTSNVLGSGSGVFKCG	650
Qy	601	IAPVPSRWYYSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	660
Db	651	IAPVPSRWYYSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	710
Qy	661	VHFQSAQISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHFHKQCFSLP	716
Db	711	VHFQSAQISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHFHKQCFSLP	766
RESULT 3			
Q5R7G7_PONPY			
ID	Q5R7G7_PONPY	PRELIMINARY;	PRT; 765 AA.
AC	Q5R7G7;		
DT	01-FEB-2005	(T-EMBLrel. 29, Created)	
DT	01-FEB-2005	(T-EMBLrel. 29, Last sequence update)	
DT	01-FEB-2005	(T-EMBLrel. 29, Last annotation update)	
DE	Hypochemical protein DKFZp469P1419.		
GN	Name=DKFZp469P1419;		
OS	Pongo pygmaeus (Orangutan).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;		
OC	Pongo.		
OX	NCBI_TaxID=9600;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Kidney;		
RG	The German cDNA Consortium;		
RA	Poustka A., Albert R., Moosmayer P., Schupp I., Wellenreuther R.,		
RA	Mewes H.W., Weil B., Amid C., Omler A., Fobo G., Han M., Wiemann S.;		
RL	Submitted (NOV-2004) to the EMBL/GenBank/DBJ databases.		
DR	EMBL: CR860150; CAH92293.1; -; mRNA.		
DR	SMR; Q5R7G7; 38-765.		
DR	GO; GO:0016020; C-membrane; IEA.		
DR	GO; GO:0004177; F-aminopeptidase activity; IEA.		
DR	GO; GO:0004274; F-dipeptidyl-peptidase IV activity; IEA.		
DR	GO; GO:0016787; F-hydrolase activity; IEA.		
DR	GO; GO:0004287; F-prolyl oligopeptidase activity; IEA.		
DR	GO; GO:0006508; P-proteolysis and peptidolysis; IEA.		
DR	InterPro; IPR002471; Pept_S9AS.		
DR	InterPro; IPR001375; Peptidase S9.		
DR	InterPro; IPR002469; Peptidase S9B.		
DR	InterPro; IPR000379; Ser esters.		
DR	Pfam; PF00930; DPPIV_N; I.		
DR	Pfam; PF00326; Peptidase S9; 1.		
DR	PROSITE; PS00708; PRO_ENDOPEP_SER; 1.		
KW	Hydrolase; Hypothetical protein; Transmembrane.		
SQ	SEQUENCE 765 AA; 88038 MW; 6F81ECD98C3D2397 CRC64;		
Query Match			
Best Local Similarity 99.3%; Score 3848; DB 2; Length 765;			
Matches 710; Conservative 2; Mismatches 4; Indels 0; Gaps 0;			
Qy	1	NTYRLKLYSLRWISDHELYLKQNNILVFNAEYGNSSVFLENSTPFGHSINDYISPD	60
Db	50	NTYRLKLYSLRWISDHELYLKQNNILVFNAEYGNSSAFLENSTPFGHSINDYISPD	109
Qy	61	GQFILLEYNVQWRHSYTSYSDIYDLNKRQLITEIRIPNNTQWTVSPVGHKLAYVWNN	120
Db	110	GQFILLEYNVQWRHSYTSYSDIYDLNKRQLITEIRIPNNTQWTVSPVGHKLAYVWNN	169
Qy	121	DIYVKLEPNLPSVRIITWTGKEDIYNGITDWWYEEVFSAYSALWWSPNGTFLAYAQFND	180
Db	170	DIYVKLEPNLPSHRIITWTGKEDIYNGITDWWYEEVFSAYSALWWSPNGTFLAYAQFND	229
Qy	181	TEVPLIEYGFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDLSVNTNATSIQITAP	240
Db	230	TEVPLIEYGFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDLSVNTNATSIQITAP	289
Qy	241	ASMLIGDHYLCDVTWATERISLQWLRRIQNYSDICDYDESSGRWNCVLVARQHIEMST	300
Db	290	ASMLIGDHYLCDVTWATERISLQWLRRIQNYSDICDYDESSGRWNCVLVARQHIEMST	349

Qy	301	TGMVGRFRSEPHFTLDGNSFYKLIISNEEGYRHCYFQIDKKDCTBITTKGTWEVIGIAL	360
Db	350	TGMVGRFRSEPHFTSDGNSFYKLIISNEEGYRHCYFQIDKKDCTBITTKGTWEVIGIAL	409
Qy	361	TSGLYLYISNEYKGMPPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	420
Db	410	TSGLYLYISNEYKGMPPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKEAKYQL	469
Qy	421	RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQML	480
Db	470	RCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWYQML	529
Qy	481	PPHFDKSKYPLLLDVIYAGCSQKADTVFRLNWTATLASTENIIIVASFGRSGSYQGDKI	540
Db	530	PPHFDKSKYPLLLDVIYAGCSQKADTVFRLNWTATLASTENIIIVASFGRSGSYQGDKI	589
Qy	541	MHAINRRLGTFFVEVDQIEAARQFSKMGFVDNKRILAIWGSYGVTSMVLGSGGVFKCG	600
Db	590	MHAINRRLGTFFVEVDQIEAARQFSKMGFVDNKRILAIWGSYGVTSMVLGSGGVFKCG	649
Qy	601	IAPVPSRWYYSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	660
Db	650	IAPVPSRWYYSVYTERYMGFLPTEDNLDHYRNSTVMSRAENFKQVEYLLIHGTADDN	709
Qy	661	VHFQSAQISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHFHKQCFSLP	716
Db	710	VHFQSAQISKALVDGVDFQAMWYTDDEHGIASTAHQHIYTHMSHFHKQCFSLP	765
RESULT 4			
DPP4_BOVIN			
ID	DPP4_BOVIN	STANDARD;	PRT; 765 AA.
AC	P81425; Q8WVG8;		
DT	15-DEC-1998	(Rel. 37, Created)	
DT	28-FEB-2003	(Rel. 41, Last sequence update)	
DT	13-SEP-2005	(Rel. 48, Last annotation update)	
DE	Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP		
DE	IV) (T-cell activation antigen CD26) (Adenosine deaminase complexing		
DE	protein) (ADCP-I) (Activation molecule 3) (ACT3) (WC10) [Contains:		
DE	Dipeptidyl peptidase 4 membrane form (Dipeptidyl peptidase IV membrane		
DE	form); Dipeptidyl peptidase 4 soluble form (Dipeptidyl peptidase IV		
DE	soluble form);		
GN	Name=DPP4; Synonyms=CD26;		
OS	Bos taurus (Bovine).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;		
OC	Bovidae; Bovinae; Bos.		
OX	NCBI_TaxID=9913;		
RN	[1]		
RP	NUCLEOTIDE SEQUENCE.		
RC	TISSUE=Lymphocyte;		
RX	MEDLINE=22067734; PubMed=12073152; DOI=10.1007/s00251-002-0456-6;		
RA	Lee S.-U., Park Y.-H., Davis W.C., Hamilton M.J., Naessens J.,		
RA	Bohach G.A.;		
RT	"Molecular characterization of bovine CD26 upregulated by a		
RT	staphylococcal superantigen.";		
RL	Immunogenetics 54:216-220(2002).		
RN	[2]		
RP	NUCLEOTIDE SEQUENCE, AND PROTEIN SEQUENCE OF 1-17.		
RC	TISSUE=Thymus;		
RX	MEDLINE=22021197; PubMed=11981836;		
RX	DOI=10.1007/s121-4141(200205)32:5<1472::AID-IMMU1472>3.0.CO;2-Q;		
RA	Gliddon D.R., Howard C.;		
RT	"CD26 is expressed on a restricted subpopulation of dendritic cells in		
RT	vivo.";		
RL	Eur. J. Immunol. 32:1472-1481(2002).		
RN	[3]		
RP	PROTEIN SEQUENCE OF 1-24.		
RC	TISSUE=T-cell;		
RX	MEDLINE=21482004; PubMed=11598101;		
RX	DOI=10.1128/IAI.69.11.7190-7193.2001;		
RA	Lee S.-U., Ferens W., Davis W.C., Hamilton M.J., Park Y.-H., Fox L.K.,		

RA Naessens J., Bobach G.A.;
 RT "Identity of activation molecule 3 on superantigen-stimulated bovine
 RL cells is CD26.";
 RN Infect. Immun. 69:7190-7193(2001).
 RN [4]
 RP PROTEIN SEQUENCE OF 537-546.
 RC TISSUE=Kidney;
 RX MEDLINE=98293306; PubMed=9628661; DOI=10.1016/S0305-0491(97)00327-1;
 RA Ben-Shooshan I., Parola A.H.;
 RT "The Cp-I subunit of adenosine deaminase complexing protein from calf
 RT kidney is identical to human, mouse, and rat dipeptidyl peptidase
 RT IV.";
 RL Comp. Biochem. Physiol. 119B:289-292(1998).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini provided that the
 CC penultimate residue is proline. Binds and regulates the activity
 CC of ADA.
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form (By similarity).
 CC -1- TISSUE SPECIFICITY: Intestinal epithelium, dendritic cells and
 CC several immune system tissues.
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
 CC by proteolytic processing (By similarity).
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 DR ENBL; AF461806; AAL67836.1; -; mRNA.
 DR ENBL; AY056834; AAL23628.1; -; mRNA.
 DR HSSP; P27487; 1PFP.
 DR SMR; P81425; 38-764.
 DR MEROPS; S09.003; -.
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser estrs.
 DR Pfam; PF00930; DPPIV N1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Aminopeptidase; direct protein sequencing; Glycoprotein; Hydrolase;
 KW Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 765
 FT Dipeptidyl peptidase 4 soluble form (By
 FT similarity).
 FT TOPO_DOM 1 6
 FT Cytoplasmic (Potential).
 FT TRANSMEM 7 29
 FT Signal-anchor for type II membrane
 FT protein (Potential).
 FT Extracellular (Potential).
 FT ACT_SITE 30 765
 FT Charge relay system (By similarity).
 FT ACT_SITE 629 629
 FT Charge relay system (By similarity).
 FT ACT_SITE 707 707
 FT Charge relay system (By similarity).
 FT CARBOHYD 739 739
 FT Charge relay system (By similarity).
 FT CARBOHYD 84 84
 FT N-linked (GLCNAC...) (By similarity).
 FT CARBOHYD 91 91
 FT N-linked (GLCNAC...) (Potential).
 FT CARBOHYD 149 149
 FT N-linked (GLCNAC...) (Potential).
 FT CARBOHYD 218 218
 FT N-linked (GLCNAC...) (By similarity).
 FT CARBOHYD 228 228
 FT N-linked (GLCNAC...) (By similarity).
 FT CARBOHYD 271 271
 FT N-linked (GLCNAC...) (Potential).
 FT CARBOHYD 280 280
 FT N-linked (GLCNAC...) (By similarity).
 FT CARBOHYD 320 320
 FT N-linked (GLCNAC...) (By similarity).
 FT CARBOHYD 392 392
 FT N-linked (GLCNAC...) (Potential).
 FT CARBOHYD 495 495
 FT N-linked (GLCNAC...) (Potential).
 FT CARBOHYD 684 684
 FT By similarity.
 FT DISULFID 384 393
 FT By similarity.
 FT DISULFID 443 446
 FT By similarity.
 FT DISULFID 453 471
 FT By similarity.

FT DISULFID 648 761 By similarity.
 SQ SEQUENCE 765 AA; 88369 MW; E32165421F43E116 CRC64;
 Query Match 91.1%; Score 3531; DB 1; Length 765;
 Best Local Similarity 89.4%; Pred. No. 1.6e-232;
 Matches 639; Conservative 40; Mismatches 36; Indels 0; Gaps 0;
 QY 1 NTYRLKLYSLRWISDHEYLKQENNLVFNAAEYGNSSVFLNENSTFDFEFGHSHINDYSISPD 60
 DB 50 NTFPMKLYNLRWSDHEYLKQENNLVFNAAEYGNSSVFLNENSTFDFEFGHSHINDYSISPD 109
 QY 61 GQFILLVYVVKWRHSYASDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 120
 DB 110 ROYILFEYVYVVKWRHSYASDIYDLNKRQLITEERIPNNTQWTVTSPVGHKLAYVWNN 169
 QY 121 DIVVKTEPNLPSVRIWTGKEDIYNGITDWWVEEVFSAYLSALWSPNCTFLAYAQFND 180
 DB 170 DIVVKTEPNLPSVRIWTGKEDIYNGITDWWVEEVFSAYLSALWSPNCTFLAYAQFND 229
 QY 181 TEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDSLSSVTNATSIIQITAP 240
 DB 230 TEVPLIEYSFYSDLSQYPKTVRVPYKAGAVNPTVKFFVNTDSLSSVTNATSIIQITAP 289
 QY 241 ASMLIGDHVLCVWTATQERISLQWRRIQNTSVMDICDYDESSGRNCLVLRQHTEMST 300
 DB 290 GSVLIGDHVLCVWTATQERISLQWRRIQNTSVMDICDYDESSGRNCLVLRQHTEMST 349
 QY 301 TGMVGRFRSEPHFTLDGNSFYKIIISNEGYRHCYFOIDKKDCTFITTKGTWVIGIEAL 360
 DB 350 TGMVGRFRSEPHFTLDGNSFYKIIISNEGYRHCYFOIDKKDCTFITTKGTWVIGIEAL 409
 QY 361 TSDYLYYIISNEYKMGPGGRNLYKIQLSYDTKYTCLSCELNPERCQYYSVSFSKEAKYQL 420
 DB 410 TSDYLYYIISNEYKMGPGGRNLYKIQLSYDTKYTCLSCELNPERCQYYSVSFSKEAKYQL 469
 QY 421 RCGPGGLPLYTHLSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFTWOMIL 480
 DB 470 RCGPGGLPLYTHLSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDPIILNETKFTWOMIL 529
 QY 481 PPHFDSKKYPLLLDYVAGCSQKADTVFLRWATVYLASTENIIIVASFGRSGSGYQGDKI 540
 DB 530 PPHFDSKKYPLLLDYVAGCSQKADTVFLRWATVYLASTENIIIVASFGRSGSGYQGDKI 589
 QY 541 MEAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIALWGSYGGYVTSMLGSGSGYVFKCG 600
 DB 590 MEAINRRLGTFEVEDQIEAARQFSKMGFVNDKRIALWGSYGGYVTSMLGSGSGYVFKCG 649
 QY 601 IAVAPVSRWEYSDSVYTERYMGLEPTEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN 660
 DB 650 IAVAPVSRWEYSDSVYTERYMGLEPTEDNLDHYNSTVMSRAENFKQVEYLLIHGTADDN 709
 QY 661 VHFQSAQISKALVDVGVDFQAMWYTDHGHGTAASHTAHOHYTHMSHFHKQCFSL 715
 DB 710 VHFQSAQISKALVDVGVDFQAMWYTDHGHGTAASHTAHOHYTHMSHFHKQCFSL 764
 RESULT 5
 ID DPP4_FELCA STANDARD; PRT; 765 AA.
 AC Q9N2I7;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
 DE IV) (T-cell activation antigen CD26) [contains: Dipeptidyl peptidase 4
 DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
 DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
 GN Name=DPP4; Synonyms=CD26;
 OS Felis silvestris catus (Cat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Felidae;
 OC Felinae; Felis.
 OX NCBI_TaxID=9685;

RN NUCLEOTIDE SEQUENCE.
 RP TISSUE=Peripheral blood;
 RX MEDLINE=20094000; PubMed=10630304; DOI=10.1007/s002510050616;
 RA Nishimura Y., Miyazawa T., Ikeda Y., Izumiya Y., Nakamura K., Sato E.,
 RA Mikami T., Takahashi E.;
 RT "Molecular cloning and sequencing of a cDNA encoding the feline T-cell
 activation antigen CD26 homologue.";
 RL Immunogenetics 50:366-368(1999).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 CC polypeptides having unsubstituted N-termini provided that the
 CC penultimate residue is proline (By similarity).
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 CC Zaa- from a polypeptide, preferentially when Yaa is Pro, provided
 CC Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer (By similarity).
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 CC soluble form (By similarity).
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (WDPP)
 CC by proteolytic processing (By similarity).
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL; AB023952; BAA92344.1; -; mRNA.
 DR HSSP; P27487; 1PFQ.
 DR SMR; Q9N217; 38-765.
 DR MEROPS; S09.003; -;
 DR InterPro; IPR002471; Pept_S9_AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estr.
 DR Pfam; PF00930; DPPIV_N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
 KW Aminopeptidase; Glycoprotein; Hydrolase; Protease; Serine protease;
 KW Signal-anchor; Transmembrane.
 FT CHAIN 1 765 Dipeptidyl peptidase 4 membrane form.
 FT CHAIN 38 765 Dipeptidyl peptidase 4 soluble form (By
 FT similarity).
 FT TOPO_DOM 1 6 Cytoplasmic (Potential).
 FT TRANSMEM 7 29 Signal-anchor for type II membrane
 FT protein (Potential).
 FT
 FT TOPO_DOM 30 765 Extracellular (Potential).
 FT ACT_SITE 629 629 Charge relay system (By similarity).
 FT ACT_SITE 707 707 Charge relay system (By similarity).
 FT ACT_SITE 739 739 Charge relay system (By similarity).
 FT CARBOHYD 84 84 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 91 91 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 149 149 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 178 178 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 228 228 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 280 280 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 320 320 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 330 330 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 331 331 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 519 519 N-linked (GlcNAc...) (By similarity).
 FT CARBOHYD 684 684 N-linked (GlcNAc...) (By similarity).
 FT DISULFID 384 393 By similarity.
 FT DISULFID 443 446 By similarity.
 FT DISULFID 453 471 By similarity.
 FT DISULFID 648 761 By similarity.
 FT SEQUENCE 765 AA; 88213 MW; 3EFC98A22B175D9 CRC64;
 Query Match 90.3%; Score 3501; DB 1; Length 765;
 Best Local Similarity 88.0%; Pred. No. 1.8e-230;
 Matches 630; Conservative 46; Mismatches 40; Indels 0; Gaps 0;
 1 NTYRLKXLSLWISDHELYKQNNILVFNAEYGNSSVFLNSPTDFEGHNSINDYSIPD 60

Db 50 NTFRVKFKYSLRWVSDHDYLYKQDNILLFNAEYGNSSVFLNSPTDFEGHNSINDYSIPD 109
 QY 61 GQFILLENNYVKQWRHSYTSYDIYDLNKRQLITERRPNNNTQWVTWSPVGHKLAYVWNN 120
 Db 110 GQFILLENNYVKQWRHSYTSYDIYDLNKRQLITBEKIPNNNTQWVTWSPVGHKLAYVWNN 169
 QY 121 DIYVKIEPNLPSYRITWTGKEDIYNGITDWYVEEVSAYSALMWSNPGTFLAYAQFND 180
 Db 170 DYYVKNENPSSSHRITWTGEEINAYNGIADWYBEEIFSAYSALMWSNPGTFLAYAQFND 229
 QY 181 TEVPLIEYSFYDESLSQYPKTVRVPYKAGAVNPVTKFFVNTDSLSSTVNTATSIIQTAP 240
 Db 230 TQVPLIEYSFYDESLSQYPMWNRIPYKAGAAVNPVTKLVKTDNLNPNNTATSVEITPP 289
 QY 241 ASMLIGDHYLCVDTWATERISQWLRRIQNYVMDICDYDESSGRWNCVLRQHIEMST 300
 Db 290 AAMLGDDYLCVDTWANEERISQWLRRIQNYVMDIRDYNNTSGKWISSAAQEHIEEMST 349
 QY 301 TGMVGRFRPSEPHEFTLDGNSFYKIIISNEEGYRHICYFOIDKDKCTFITKGTWEVIGIEAL 360
 Db 350 TGMVGRFRPAEPHEFTSDGRNFYKIIISNBDGYKHICRFQIDKKDCTFITKGAWEVIGIEAL 409
 QY 361 TSDYLYIISNEYKMGPGGRNLYKIQLSYTKVTCLSCELNPERCOVYSVSFSEAKYQYL 420
 Db 410 TTDYLYIISNEYKMGPGGRNLYKIQLDYTKVACLSCELKPERCOVYSVSFSEAKYQYL 469
 QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLPDIFILNETKFWYQML 480
 Db 470 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQEVQMPSSKKLPDIFILNETKFWYQML 529
 QY 481 PPHFDKSKYPLLLDYAGPCSQKADTVFRLNWTYLASTENIIVASDFGRSGYQGDKI 540
 Db 530 PPHFDTSKKYPLLLDYAGPCSQKADAFRLNWTYLASTENIIVASDFGRSGYQGDKI 589
 QY 541 MHAINRRLGTPEVEDQIEAARQFSKMGFVDNKRKRIAWGWSYGGYVTSVMVLSGSGYVFKCG 600
 Db 590 MHAVNRRLGTPEVEDQIEAARQFSKMGFVDNKRKRIAWGWSYGGYVTSVMVLSGSGYVFKCG 649
 QY 601 IAVAPVSRWEYYSVYTRYMGLPTPEDNLDHNRNSTMVSRANFKQVEYLLIHGTADDN 660
 Db 650 IAVAPVSRWEYYSVYTRYMGLPTPDNLNLDYKNSVMSRAENFKQVEYLLIHGTADDN 709
 QY 661 VHFQSAQISKALVDGVDFQAMWYTDHGHGTASSTAHQHIYTHMSHFHKQCFSLP 716
 Db 710 VHFQSAQISKALVDAGVDFQAMWYTDHGHGTASSTAHQHIYTHMSHFHKQCFSLP 765
 RESULT 6
 DPP4_PIG
 ID DPP4_PIG STANDARD; PRT; 766 AA.
 AC P22411; Q866G3;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 25-OCT-2004 (Rel. 45, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Dipeptidyl peptidase 4 (EC 3.4.14.5) (Dipeptidyl peptidase IV) (DPP
 DE IV) (T-cell activation antigen CD26) [Contains: Dipeptidyl peptidase 4
 DE membrane form (Dipeptidyl peptidase IV membrane form); Dipeptidyl
 DE peptidase 4 soluble form (Dipeptidyl peptidase IV soluble form)].
 GN Name=DPP4; Synonyms=CD26;
 OS Sus scrofa (Pig)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Suina; Suidae;
 OC Sus.
 ON NCBI_TaxID=9823;
 RN [1]
 RP NUCLEOTIDE SEQUENCE, AND FUNCTION.
 RC TISSUE=Kidney;
 RX PubMed=14719797; DOI=10.1515/BC.2003.172;
 RA Baer J., Weber A., Hoffmann T., Stork J., Wermann M., Wagner L.,
 RA Aust S., Gehartz B., Demuth H.-U.;
 RT "Characterisation of human dipeptidyl peptidase IV expressed in Pichia
 RT pastoris. A structural and mechanistic comparison between the

recombinant human and the purified porcine enzyme.";
 [2]
 RN BIOL. Chem. 384:1553-1563(2003).
 RP NUCLEOTIDE SEQUENCE OF 2-67.
 RC TISSUE=Kidney;
 RX MEDLINE=94093209; PubMed=7903569;
 RA Thomsen P.D., Qvist H., Marklund L., Andersson L., Sjostrom H.,
 Noren O.;
 RT "Assignment of the dipeptidylpeptidase IV (DPP4) gene to pig
 chromosome 15q21.";
 RN Mamm. Genome 4:604-607(1993).
 RP PROTEIN SEQUENCE OF 38-71.
 RC TISSUE=Kidney;
 RX MEDLINE=91273813; PubMed=1675855;
 RA Seidl R., Mann K., Schaeffer W.;
 RT "N-terminal amino-acid sequence of pig kidney dipeptidyl peptidase IV
 solubilized by autolysis.";
 RN BIOL. Chem. Hoppe-Seyler 372:213-214(1991).
 RP X-RAY CRYSTALLOGRAPHY (1.8 ANGSTROMS) OF 39-766, N-GLYCOSYLATION
 SITES, AND HOMODIMERIZATION.
 RX PubMed=12690074; DOI=10.1073/pnas.0230620100;
 RA Engel M., Hoffmann T., Wagner L., Wermann M., Heiser U.,
 Kieffersauer R., Huber R., Bode W., Demuth H.-U., Brandstetter H.;
 RT "The crystal structure of dipeptidyl peptidase IV (CD26) reveals its
 functional regulation and enzymatic mechanism.";
 RN Proc. Natl. Acad. Sci. U.S.A. 100:5063-5068(2003).
 CC -1- FUNCTION: Removes N-terminal dipeptides sequentially from
 polypeptides having unsubstituted N-termini provided that the
 penultimate residue is proline.
 CC -1- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
 Zaa, from a polypeptide, preferentially when Yaa is Pro, provided
 Zaa is neither Pro nor hydroxyproline.
 CC -1- SUBUNIT: Homodimer.
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
 soluble form.
 CC -1- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
 by proteolytic processing.
 CC -1- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use as long as its content is in no way modified and this statement is not
 removed.

 DR ENBL; AY198323; AA043404.1; -; mRNA.
 DR ENBL; X73276; CAAS1717.1; -; mRNA.
 DR PIR; I47134; I47134.
 DR PDB; 1ORV; X-ray; A/B/C/D=39-766.
 DR PDB; 1ORW; X-ray; A/B/C/D=39-766.
 DR SNR; P22411; 39-766.
 DR MEROPS; S09.003; -.
 DR InterPro; IPR002471; Pept S9 AS.
 DR InterPro; IPR001375; Peptidase_S9.
 DR InterPro; IPR002469; Peptidase_S9B.
 DR InterPro; IPR000379; Ser_estr.
 DR Pfam; PF00930; DPPIV N; 1.
 DR Pfam; PF00326; Peptidase_S9; 1.
 DR PROSITE; PS00708; PRO_ENDOPEP_SBR; 1.
 KW 3D-structure; Aminopeptidase; Direct protein sequencing; Glycoprotein;
 Hydrolase; Protease; Serine protease; Signal-anchor; Transmembrane.
 FT CHAIN 1 766 Dipeptidyl peptidase 4 membrane form.
 FT CHAIN 38 766 Dipeptidyl peptidase 4 soluble form.
 FT TOPO_DOM 1 6 Cytoplasmic (Potential).
 FT TRANSMEM 7 27 Signal-anchor for type II membrane
 protein (Potential).
 FT TOPO_DOM 28 766 Extracellular (Potential).
 FT ACT_SITE 630 766 Charge relay system (By similarity).
 FT ACT_SITE 708 766 Charge relay system (By similarity).
 FT ACT_SITE 740 766 Charge relay system (By similarity).
 FT ACT_SITE 85 85 N-linked (GlcNac...).
 FT CARBOHYD

FT	CARBOHYD	92	92	N-linked (GlcNac...)	
FT	CARBOHYD	150	150	N-linked (GlcNac...)	(By similarity).
FT	CARBOHYD	179	179	N-linked (GlcNac...)	(Potential).
FT	CARBOHYD	219	219	N-linked (GlcNac...)	(By similarity).
FT	CARBOHYD	229	229	N-linked (GlcNac...)	
FT	CARBOHYD	279	279	N-linked (GlcNac...)	
FT	CARBOHYD	321	321	N-linked (GlcNac...)	
FT	CARBOHYD	393	393	N-linked (GlcNac...)	(Potential).
FT	CARBOHYD	685	685	N-linked (GlcNac...)	
FT	DISULFID	385	394		
FT	DISULFID	444	447		
FT	DISULFID	454	472		
FT	DISULFID	649	762		
FT	CONFLICT	32	32	Missing (in Ref. 2).	
SQ	SEQUENCE	766 AA;	88242 MW;	8800D520BAEA856D CRC64;	
Query Match 90.0%; Score 3489; DB 1; Length 766;					
Best Local Similarity 88.1%; Pred. NO. 1.2e-229;					
Matches 631; Conservative 45; Mismatches 40; Indels 0; Gaps 0;					
QY	1	NTVRLKLSLRWISDRHEYLKQENNLVFNNAEYGNSSVFLNSTFDFBGHSINDYSISPD	60		
DB	51	STFRVAFYTLQWISDRHEYLKQENNLVFNNAEYGNSSVFLNSTFDFBGHSINDYSISPD	110		
QY	61	GQFILLVYVVKQWRHSYTSYDIYDLNKRQLITEBIPNNTQWTVTSPVGHKLAYVWNN	120		
DB	111	RQFILLVYVVKQWRHSYTSYDIYDLNKRQLITEBIPNNTQWTVTSPVGHKLAYVWNN	170		
QY	121	DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSALMWSNPGTFLAYAQFND	180		
DB	171	DIYVKLEPNLPSYRITWTGKEDIYNGITDWTVEEVFSAYSALMWSNPGTFLAYAQFND	230		
QY	181	TEVPLEYFSYDESQYKPTVRVPYKAGVNPVTKFVWNTDLSVSNATSIQITAP	240		
DB	231	TEVPLEYFSYDESQYKPTVRVPYKAGVNPVTKFVWNTDLSVSNATSIQITAP	290		
QY	241	ASMLGDHGLCDVTWATQERISLQWLRRIQNTSVMDICDYDESGRWNCVLARQHIEMST	300		
DB	291	ASVLGDHGLCDVTWATQERISLQWLRRIQNTSVMDICDYDESGRWNCVLARQHIEMST	350		
QY	301	TGWVGRFRSEPHFTLDGNSFYKLSNBEQYRHIQYFQDKDKDCTTITGKTWEVIGIEAL	360		
DB	351	TGWVGRFRSEPHFTLDGNSFYKLSNBEQYRHIQYFQDKDKDCTTITGKTWEVIGIEAL	410		
QY	361	TSDDLVIYIENEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQL	420		
DB	411	TSDDLVIYIENEYKMGPGGRNLKYQLSDYTKVTCLSCELNPERCOYYSVSFSKEAYQL	470		
QY	421	RCSGPGPLTYTLHSSVNDKGLRVLEDNSALDKMLQVQMPKSKLDPIILNETKFWYQML	480		
DB	471	RCFGPGPLTYTLHSSVNDKGLRVLEDNSALDKMLQVQMPKSKLDPIILNETKFWYQML	530		
QY	481	PHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIIVASFDGRSGYQGDKI	540		
DB	531	PHFDKSKYPLLLDVYAGCSQKADTVFRLNWTYLASTENIIIVASFDGRSGYQGDKI	590		
QY	541	MEAINRLCTFEVEDOIEAARQFSKMGFVDNKRKRIATWGHSGYGYVTSWVLGSSGVFKC	600		
DB	591	MEAINRLCTFEVEDOIEAARQFSKMGFVDNKRKRIATWGHSGYGYVTSWVLGSSGVFKC	650		
QY	601	IIVAPVSRWEYDVSVYTERYMGLPEDNLDHNRSTVMSRAENPKQVEYLHGTADDN	660		
DB	651	IIVAPVSRWEYDVSVYTERYMGLPEDNLDHNRSTVMSRAENPKQVEYLHGTADDN	710		
QY	661	VHFQSAQISKALVDVGVDFOAMWYTDDEHGIASTAHQHIYTHMSHFHKQCFSLP	716		
DB	711	VHFQSAQISKALVDVGVDFOAMWYTDDEHGIASTAHQHIYTHMSHFHKQCFSLP	766		
RESULT 7					
DPP4 RAT					
ID	DPP4 RAT				
AC	P14740;				
	STANDARD;				
	PRT;				
	767 AA.				

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Heiton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Glickman J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalaka U., Smalilus D.B.,
RA Schnurch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RL and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [6]
RP PROTEIN SEQUENCE OF 1-20.
RX MEDLINE-91302787; PubMed-1712807;
RA Vivier I., Marguet D.A., Naquet P., Bonicel J., Black D., Li C.X.-Y.,
RA Bernard A.-M., Gorvel J.-P., Pierres M.;
RA "Evidence that thymocyte-activating molecule is mouse CD26 (dipeptidyl
RT peptidase IV).";
RT J. Immunol. 147:447-454(1991).
CC -I- FUNCTION: Removes N-terminal dipeptides sequentially from
CC polypeptides having unsubstituted N-termini provided that the
CC penultimate residue is proline.
CC -I- CATALYTIC ACTIVITY: Release of an N-terminal dipeptide, Xaa-Yaa-|-
CC Zaa-, from a polypeptide, preferentially when Yaa is Pro, provided
CC Zaa is neither Pro nor hydroxyproline.
CC -I- SUBUNIT: Homodimer.
CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Also exists in a
CC soluble form.
CC -I- PTM: The soluble form (SDPP) derives from the membrane form (MDPP)
CC by proteolytic processing.
CC -I- SIMILARITY: Belongs to the peptidase S9B family. DPPIV subfamily.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL; X58384; CAA41274.1; -; Genomic DNA.
DR EMBL; U12620; AAA82213.1; -; Genomic DNA.
DR EMBL; U12599; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12600; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12601; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12602; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12603; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12604; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12605; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12606; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12607; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12608; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12609; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12610; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12611; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12612; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12613; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12614; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12615; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12616; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12617; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12618; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; U12619; AAA82213.1; JOINED; Genomic DNA.
DR EMBL; AK085370; BAC39434.1; -; mRNA.
DR EMBL; BC022183; AAH22183.1; -; mRNA.
DR HSSP; P27487; 1PFO.
DR SMR; P28843; 37-759.
DR MEROPS; S09.003; -.
DR Ensembl; ENSMUSG00000035000; Mus musculus.
DR MGI; MG1:94919; Dpp4.
DR GO; GO:0016021; C:integral to membrane; TAS.
DR GO; GO:0046581; C:intercellular canalliculus; IDA.
DR InterPro; IPR002471; Pept S9 AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.

DR InterPro; IPR000379; Ser.estr.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Aminopeptidase; Direct protein sequencing; Glycoprotein; Hydrolase;
KW Protease; Serine protease; Signal-anchor; Transmembrane.
FT CHAIN 1 760 Dipeptidyl peptidase 4 membrane form.
FT CHAIN 37 760 Dipeptidyl peptidase 4 soluble form (By
FT similarity).
FT TOPO_DOM 1 6 Cytoplasmic (Potential).
FT TRANSMEM 7 28 Signal-anchor for type II membrane
FT protein (Potential).
FT TOPO_DOM 29 760 Extracellular (Potential).
FT COMEBIAS 295 466 Cys-rich.
FT ACT_SITE 624 624 Charge relay system (By similarity).
FT ACT_SITE 702 702 Charge relay system (By similarity).
FT ACT_SITE 734 734 Charge relay system (By similarity).
FT CARBOHYD 83 83 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 90 90 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 213 213 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 223 223 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 315 315 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 328 328 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 514 514 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 679 679 N-linked (GlcNAc...) (By similarity).
FT DISULFID 379 388 By similarity.
FT DISULFID 438 441 By similarity.
FT DISULFID 448 466 By similarity.
FT DISULFID 643 756 By similarity.
SQ SEQUENCE 760 AA; 87437 MW; ASF644B46B43DF8 CRC64;
Query Match 86.1%; Score 3340; DB 1; Length 760;
Best Local Similarity 84.9%; Pred. No. 1.8e-219;
Matches 607; Conservative 55; Mismatches 49; Indels 4; Gaps 1;
QY 1 NTRYRLKSLRWISDHEYLKQENNLVFNABYGNSSVFLENSTDFEGHNSINDYSISDP 60
DB 49 STFRVKSLSLWVSDFEYLKQENLLLLNAEHGSSIFLENSTFESFGY----HSVSDP 104
QY 61 GQFILLENNYKQWRHSYASDYIDLNKRQLITEERIPNNQWYTWGSHGLAYVMNN 120
DB 105 RLFLVLENNYKQWRHSYASVNIYDNKRQLITEEKIPNNQWITWSPGSHGLAYVMNN 164
QY 121 DIYVKLENNLPSYRITWTCKEDIIYNGITDWYEEVESAYSALWSPNGTFLAYAQND 180
DB 165 DIYVKVEPLPSHRITSTGEENVYNGITDWYEEVEFGVSAWSPNNTFLAYAQFND 224
QY 181 TSVPLIEYSFYSDESLSQYPKTVRPYPKAGAVNTPVKPVNNTSLSSVTNATSIQITAP 240
DB 225 TGVPLIEYSFYSDESLSQYPKTVWIPYKAGAVNTPVKPVNNTSLSSSSSAAPQIPAP 284
QY 241 ASMLIGDHYLCDVTWATERISLOWLRRIQNYVMDICDYDESSGRWNCVLAQRHIEMST 300
DB 285 ASVARGDHYLCDVWVWATERISLOWLRRIQNYVMAICDYDKINLTWNCPSQQRHVEMST 344
QY 301 TGWGRFRPSEPHFTLDGNSFYKIIISNEGYRHICYFQIDKDKCTFITKGTWEVIGIEAL 360
DB 345 TGWGRFRPAEPHFTSDGSSFYKIIISDKDGYKHCHFPKDKDCTFITKGTWEVIGIEAL 404
QY 361 TSDYLYIYSNEYKMGPGGRNLYKIQLSDYTKVTCVLSCELNPERCOVYSFSEKARYQL 420
DB 405 TSDYLYIYSNQYKPNPGGRNLYKIQLTDHTVNVKCLSDLNPERCOVYAVSFSEKARYQL 464
QY 421 RCGSGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFIILNETKFWQMIL 480
DB 465 GCMGPGPLPLYTLHRTDHLKRLVLEDNSALDKMLQDVQMPSSKKLDFIILNETKFWQMIL 524
QY 481 PPHFDKSKYIPLLLDVYAGPCSQKADTVFRLNWTATYLASTENIIVASPDGRSGVQGDKI 540
DB 525 PPHFDKSKYIPLLLDVYAGPCSQKADAFRLNWTATYLASTENIIVASPDGRSGVQGDKI 584
QY 541 MHAINRRLLGTPEVEDQIEAARQFSKMGFVDNKRITAIWGSVGYGYVTSMVLGSGSVGFKCG 600

```

Db 585 MHAINRLGTLVEDEQIEAARQFVKMGFVDSKRVAINWGSYGGVYVTSNVLGSGSVFKCG 644
Qy 601 IAVAPVRWEYDVSYYTERYMGTPEDNLDHYRNSVTMSRAENPKQVQVYLLIHGTADDN 660
Db 645 IAVAPVRWEYDVSYYTERYMGTPEDNLDHYRNSVTMSRAENPKQVQVYLLIHGTADDN 704
Qy 661 VHQFQSAQISKALVDVGVDFQAMWYTDHEDHGIASTAHQIYTHMSHFIIKQCFSL 715
Db 705 VHQFQSAQISKALVDAGVDFQAMWYTDHEDHGIASTAHQIYTHMSHFIIKQCFSL 759

RESULT 9
Q75S83 AGKHB
ID Q75S83 AGKHB PRELIMINARY; PRT; 751 AA.
AC Q75S83;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Dipeptidylpeptidase 4a (EC 3.4.14.5).
GN Name=DP4a;
OS Agkistrodon haly's brevicaudus (Korean slamosa snake) (Gloydus haly's brevicaudus);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=259325;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Venom gland;
RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB158224; BAD06332.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9 AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV N; 1.
DR PROSITE; PS00708; PRO-ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 751 AA; 86127 MW; 6A3F2D829BE97D4 CRC64;

Query Match 64.2%; Score 2490.5; DB 2; Length 751;
Best Local Similarity 62.9%; Pred. No. 2e-161;
Matches 449; Conservative 110; Mismatches 144; Indels 11; Gaps 5;

Qy 3 YRLKLSLRWISDHEVLYKQENILVFNABYGNSSVFLNSTDFEGHSINDYSISPDGQ 62
Db 49 FOYKSYNLRWMSGHEVYVYTNQNNVLLYNDRESIVLSNDTLDSEFSS--QAILSPDRK 106

63 FILLEYNVKWRHSYVSTASDIYDLNKRQLITRERIPNNTQWTSVPVGHKLAVVWVNDI 122
Db 107 FALQVSYEKWRHSYVSTASDIYDLNKRQLITRERIPNNTQWTSVPVGHKLAVVWVNDI 166

123 YVKIEPNLPSYRITGTGKEDIIYNGITDVIYEEVFSAYSGALWMSNGTFLAYAQNDTE 182
Db 167 YVKATPNASPVQITENGAEKNILNGLADWYEEEMFGTHSALWSPNGRFLAEINDTE 226

183 VPLIEYSFYDESLOVPTKTVRVPYKAGAVNPTVKFPVNVNTDLSLSSVYTNATSIQTAPAS 242
Db 227 VPWMEYSFYSEDLTQVPTKTIKIPYKAGAVNPTVKFPVNVNTDLSLSSVYTNATSIQTAPAS 282

243 MLIGDHYLCVDTWATQERISLOWLRIONVYVMDICDYDESSGRWNCVLAQHIENSTTC 302
Db 283 IISGDHYLSAVTWVTERICQWLRIONVYVMDICDYDESSGRWNCVLAQHIENSTTC 339

303 WVGFRFRPSEPHFTLDGNSFYKIIISNBEYGRHICVFQIDKDKDCTFITKGTWETVIGIEALTS 362

```

```

Db 340 WVGFRFQSEPYFTSDKISYYRIISDSGYKHHTDSAGK-VKPIITSGKWEVISAVTN 398
Qy 363 DYLYYSINSEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVFSKEAKYQJRC 422
Db 399 NSLYFISNEPEGRPGGRHLYKVLDLKNLKKICITCNSEKACQYFSVFSFSDSYKLNLC 458
Qy 423 SGPLPLTYTLHSSVNDKGLRVLBENSALDKMLQNVQMPKLDPIILNETKFWQMLPP 482
Db 459 YGPDLPYFTLQNSITDKAIKTLNENNLKVLKEIQMPCRLSNITLHGQTYWQMLPP 518
Qy 483 HFDKSKYKPLLLDVYAGPCSQKADTVFRLNWTYLASTENIIVASPDGRSGYQDKIMH 542
Db 519 NFDKSKYKPLLLDVYAGPCSQKADTVFRLNWTYLASTENIIVASPDGRSGYQDKIMH 578
Qy 543 AINRRLGTPEVEDQIEAARQFVKMGFVDSKRVAINWGSYGGVYVTSNVLGSGSVFKCGIA 602
Db 579 AIYRRLGTPEVEDQIEAARQFVKMGFVDSKRVAINWGSYGGVYVTSNVLGSGSVFKCGIA 638
Qy 603 VAPVRWEYDVSYYTERYMGTPEDNLDHYRNSVTMSRAENPKQVQVYLLIHGTADDNVH 662
Db 639 VAPVRWEYDVSYYTERYMGTPEDNLDHYRNSVTMSRAENPKQVQVYLLIHGTADDNVH 698
Qy 663 FQFSAQISKALVDVGVDFQAMWYTDHEDHGIASTAHQIYTHMSHFIIKQCFSLP 716
Db 699 FQFSAQISKALVDVGVDFQAMWYTDHEDHGIASTAHQIYTHMSHFIIKQCFSLP 751

RESULT 10
Q75S82 AGKHB
ID Q75S82 AGKHB PRELIMINARY; PRT; 751 AA.
AC Q75S82;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Dipeptidylpeptidase 4b (EC 3.4.14.5).
GN Name=DP4b;
OS Agkistrodon haly's brevicaudus (Korean slamosa snake) (Gloydus haly's brevicaudus);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Lepidosauria; Squamata; Scleroglossa; Serpentes; Colubroidae;
OC Viperidae; Crotalinae; Gloydus.
OX NCBI_TaxID=259325;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC TISSUE=Venom gland;
RA Ogawa Y., Yanoshita R., Murayama N., Higuchi S., Samejima Y.;
RL Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB158225; BAD06333.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9 AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR Pfam; PF00930; DPPIV N; 1.
DR PROSITE; PS00708; PRO-ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 751 AA; 86155 MW; DAF219B4FEE7629A CRC64;

Query Match 64.2%; Score 2489.5; DB 2; Length 751;
Best Local Similarity 62.9%; Pred. No. 2.4e-161;
Matches 449; Conservative 110; Mismatches 144; Indels 11; Gaps 5;

Qy 3 YRLKLSLRWISDHEVLYKQENILVFNABYGNSSVFLNSTDFEGHSINDYSISPDGQ 62
Db 49 FOYKSYNLRWMSGHEVYVYTNQNNVLLYNDRESIVLSNDTLDSEFSS--QAILSPDRK 106

```

```
QY 63 FILLEYVYKQWRHSYTSASDIYDLNKRQLITEIRPNNTQWTVSPVGHKLAYVWNNDI 122
Db 107 FALQSYSEKVRHRSYSTASHIYDLNRRTKITENPLPTNIQIYSWSVGHKLAYVRNV 166
QY 123 YKIEBNLPSYRITWTGKEDIIYNGITDWYEEVFSAYSALWWSPNGTFLAYAOFNDETE 182
Db 167 YVKATNAPSVOQITENGAENKILGLADWYEEEMFGTHSALWWSPNGRFLAFAEINDTE 226
QY 183 VPLIEVSFSDLSQYKTVRVYPYKAGAVNPVFFVNTDLSLSTVNTATSIIQITAPAS 242
Db 227 VPVMEYSFSEDLTQYPKTIKIPYKAGAINPTIRLFVLDI----SLSPKNTSEIIVAPSS 282
QY 243 MLIGDHYLADVNTWATQERISLOWLRIONYSVMDICDYDESSGRWNCVLAHQHIESTTG 302
Db 283 IISGDHYLSVVTWVTDERICVQWLRRIQNFSLVITICDY---SGAMWCPKREHLEESKGTG 339
QY 303 WYGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFOIDKKDCTFITKGTWEVIGIBALTS 362
Db 340 WYGRFQPSPEYFTSDKISYRIISDSEGYKHIIHYTDSACK-VKPIITSGKWEVISISAVTN 398
QY 363 DYLYISNEYKMGPGGRNLYKIQLSDYTKVTCLSCELNPERCQYYSVSFSKBAKYYQLRC 422
Db 399 NSLYFISNEPEGRPGGRHLYKVDLKNLKKICITCNSKBEACQYFSVSFSTDSRYYYKLC 458
QY 423 SGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFFILNETKWEYQMLPP 482
Db 459 YGDLPLPYTLQNSITDKAIKLTEDNNLNKNVLKEIQMPCRKLSNITLHGQTYWYQMLPP 518
QY 483 HFDKSKYKPLLLDVYAGPQSQAADTVFRLNWTATYLASTENIIIVASPDGSGGYQSGDKIMH 542
Db 519 NFDESKYKPLLLDVYAGPQSQAADAFAFRINWSTYLLASSEGIIVASPDGSGGYQSGDKILH 578
QY 543 AINRRLGTVEVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSVWLGSQGVFKCGTA 602
Db 579 AIYRRLGTVEVEDQIEAARLFSEMSFVDKRIAIWGSYGGYVTSVWLGSQGVFKCGTA 638
QY 603 VAPVSRWEYVDSYVTERYVGLPTPEDNLDHYRNSVTMSRAENPKQVEYLLIHGTADDNVH 662
Db 639 VAPVSRWQYDSYVTERYVGLPKPNLNFYENSTVWAKNFRVTDYLLIHGTADDNVH 698
QY 663 FQQAQISKALVDVGVDFQAMWYTDDEHGIASSTAHOHYTHMSHPFKOCFSLP 716
Db 699 FQQAQISKALVDVGVDFQAMWYTDKDHGIGGH-AHSHIYQHSHPMKOCFKLP 751
```

RESULT 11

```
Q52181 CHICK
ID Q52181_CHICK PRELIMINARY; PRT; 759 AA.
AC Q52181_
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=RCJMB04.29g21;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CB; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezubov Y., Zaim J.,
RA Fiedler P., Kutter S., Blagodatski A., Kostovska D., Koter M.,
RA Flachy J., Carninci P., Hayashizaki Y., Buerstedde J.M.;
RT "Full-length cDNAs from chicken bursal lymphocytes to facilitate
RT gene function analysis".
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ270903; CAG32562.1; -; mRNA.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
```

```
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; F:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR00379; Ser_estra_.
DR Pfam; PF00930; DPPIV_N; I.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Hypothetical protein; Transmembrane.
SQ SEQUENCE 759 AA; 86343 MW; B9A0C38CCB8644AC CRC64;
```

Query Match

Best Local Similarity 63.9%; Score 2478; DB 2; Length 759;

Matches 466; Conservative 92; Mismatches 147; Indels 16; Gaps 9;

```
QY 1 NTRYLKLXSLRWISDHEYLKQEN-NILVFNAEYGNSSVFLSENSTFDEFGHSINDYSISP 59
Db 50 NDYVYVYKTNLQWISGNQYLHETSNGNLLRFDAETGTSSVLLNTTISI--HEATTAILSP 107
QY 60 DQGFILLEYVYKQWRHSYTSASDIYDLNKRQLITEIRPNNTQWTVSPVGHKLAYVWN 119
Db 108 DQRFALLOQYKYEKLWRHSYTSASDIYDFNTSSILDALLPNDTQYISWSVGHKLAYVWN 167
QY 120 NDIYVYKIEPNLPSYRITWTGKEDIIYNGITDWYEEVFSAYSALWWSPNGTFLAYAOEN 179
Db 168 NNIIYKASPTAAPVQITNSGENKIFNGIPDWYEEEMFGSHALWWSPNGNPFVAYAAEN 227
QY 180 DTEVPLIEYSFSDLSQYKTVRVYPYKAGAVNPVFFVNTDLSLSTVNTATSIIQITA 239
Db 228 DTEVPEVYSFSEDLTQYPKTIRIPYKAGAKNPTVKEFIVD---IQMLPDFNSTEISP 284
QY 240 PASMLIGHYLCVDTWATQERISLOWLRIONYSVMDICDYDESSGRWNCVLAHQH---I 296
Db 285 PAEIKSGDHYLSVVTWVTDERICLQWLRRIQNSVLTICDPESATGNMTC--HRKNPFWK 342
QY 297 ENSTTGWVGRFRPSPHFTLDGNSFYKIIISNEEGYRHICYFOIDKKDCTFITKGTWEVIG 356
Db 343 KVQLAGN--QISAICPLFAPDNTTYIKVFSNTEGYKHHIYNGTEAPVP-ITEGRKEVIS 399
QY 357 IEALTSDYLYISNEYKMGPGGRNLYKIQLSDYTKVT-CLSCELNPERCQYYSVSFSKEA 415
Db 400 IAAVTKYFLYIYSNQNGEMPGGRNLYKMLLESSPKSTQCVSCDLNQERQYIYSASFSDKA 459
QY 416 KYQYLRCSGPGPLPLYTLHSSVNDKGLRVLEDNSALDKMLQNVQMPKSKLDFFILNETKFW 475
Db 460 QYQYLNCLGPGLPMTSLHRSSDDQVLRVLENNTELENSLKDIQMPSKKLGSIITVGGYNLW 519
QY 476 YQMILPPHFDKSKYKPLLLDVYAGPQSQAADTVFRLNWTATYLASTENIIIVASPDGSGGY 535
Db 520 YQMILPPHLDSSKYPPLLELVYAGPQSQAADTVFRLNWTATYLASTEIIIVASPDGSGGY 579
QY 536 QGDGKIIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSVWLGSQGS 595
Db 580 QGDGKIIMHAINRRLGTFFVEDQIEAARQFSKMGFVNDKRIAIWGSYGGYVTSVWLGSQGS 639
QY 596 VPKCGIAVAPVSRWYVDSYVTERYVGLPTPEDNLDHYRNSVTMSRAENPKQVEYLLIHG 655
Db 640 VPKCGIAVAPVSRWYVDSYVTERYVGLPTPEDNLDHYRNSVTMSRAENPKQVEYLLIHG 699
QY 656 TADDNVHFQOASQAIKALVDVGVDFQAMWYTDDEHGIASSTAHOHYTHMSHPFKOCFSL 715
Db 700 TADDNVHFQOASQAIKALVDVGVDFQAMWYTDKDHGI--SGQAHKHIIYTHMSHPFKOCFSL 758
QY 716 P 716
Db 759 P 759
```

RESULT 12

```
Q641D6 XENLA
ID Q641D6_XENLA PRELIMINARY; PRT; 737 AA.
AC Q641D6_
```



```
Db 462 YALVCGPGPISTLHGRDQEIKILEENKELENALKNLQLPKBEIKKLEVEITLWYK 521
Qy 478 MILPPHDKSKYPLLLDVTYAGCSQKADTVFRLNWTATYLASTENIIVASFDCRGSGYQG 537
Db 522 MILPPQFDRSKYPLLIQVYGGPCSVSRVFAVNWISYLASKGWIADVGRGTAFQG 581
Qy 538 DKIMHAINRLGTVEVEDQIEAARQFSKMGFVGNKRIAIWGSYGGVYVSMVLGSGGVF 597
Db 582 DKLLYAVYRKLGVYEVEDQITAVRKFTMGFIDFKRIAIWGSYGGVYVSSALASGTGLF 641
Qy 598 KCGIAPVSRWEYSDVSYTYRYMGLPTPDNDLHYRNSVTMSRAENFKQVEYLLIHGTA 657
Db 642 KCGIAPVSSWEYASVYTERFMGLPTKODNLEHYKNSTVMARAEYFRNVVDYLLIHGTA 701
Qy 658 DDNVHFOQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSL 715
Db 702 DDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGL-SGLSTNHLTYTHMTFLKQCFSL 758

RESULT 15
Q8R492 RAT PRELIMINARY; PRT; 761 AA.
AC Q8R492
DT 01-JUN-2002 (TReMBLrel. 21, Created)
DT 01-JUN-2002 (TReMBLrel. 21, Last sequence update)
DE 01-MAR-2004 (TReMBLrel. 26, Last annotation update)
DE Fibroblast activation protein alpha subunit.
GN Names-Pap;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
OC Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Zuckerswerdt K., Park J.E.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF493782; AAM11677.1; -; mRNA.
DR HSSP; P27487; 1PQF.
DR MEROPS; S09.007; -.
DR Ensembl; ENSRNOG00000005679; Rattus norvegicus.
DR RGD; 621253; Pap.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0004177; F:aminopeptidase activity; IEA.
DR GO; GO:0004274; F:dipeptidyl-peptidase IV activity; IEA.
DR GO; GO:0016787; F:hydrolase activity; IEA.
DR GO; GO:0004287; F:prolyl oligopeptidase activity; IEA.
DR GO; GO:0006508; P:proteolysis and peptidolysis; IEA.
DR InterPro; IPR002471; Pept_S9_AS.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR002469; Peptidase_S9B.
DR InterPro; IPR000379; Ser_estra.
DR Pfam; PF00930; DPPIV_N; 1.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00708; PRO_ENDOPEP_SER; 1.
KW Hydrolase; Transmembrane.
SQ SEQUENCE 761 AA; 87843 MW; BB214BBB4BB14F2C CRC64;

Query Match 55.2%; Score 2141.5; DB 2; Length 761;
Best Local Similarity 54.0%; Pred. No. 1.5e-137;
Matches 388; Conservative 128; Mismatches 189; Indels 13; Gaps 6;

Qy 2 TYRLKLYSLRWISDHEVLYKQB--NNILVFNAEYGNSSVFLNSTPDEFHGSIN--DYSIS 58
Db 51 TFSYKTYFPNWISQEYVHQSEDDNIVFYNIETRESYIILNSMT----KSVNATDGLS 106
Qy 59 PDQQFILLEYNVYKQWRHSYASYDIYDLNKRQLITEERIPNNNTQWTVWSPVGHKLAYVW 118
Db 107 PDRQFVLESYDSKLVNRYSTATYIYDLQNGEFVRGYELPRPIQYLWSPVGSKLAYVY 166
Qy 119 NNDIYKIBPNLPSYRITWTGKEDIYNGITDWWYEEVFSAWSPNGTFLAYAOFP 178
Db 167 QNNIYLKQRPDPPFOITVTGRENIFNGITPDWVYEEEMLA TKYALWSPDGKYLAYVEF 226
```

```
Qy 179 NDETEVLIIBSYFSDLESLOVPTKTRVPYPKAGAVNPTVKFFVVTNDSLSSVTNATSIQIT 238
Db 227 NDSIDIPIIAYSYYGDG--QYPRITINIPYKAGAKNPIVRVFIVDITYPHHV---GRIEVP 281
Qy 239 APASMLIGHYLCDVTWATQERISLQWLRIQNYVMDICDYDESSGRWNCLVAROHIEM 298
Db 282 VPEMIASSDYFTWLTWVTNERVCLQWLKRVQNVSVLSICDPREDHAMDCPKNQBIIEE 341
Qy 299 STTGWVGRPRPSEPHTLDGNSFYKLIISNEEGVYRHCYQIDKKDCTFTFKGTWVEVIGIE 358
Db 342 SRTGWAGGFVSTPAFSAQAAASYIKFSDKDGKYKHIIKDTVENAIQITSGKWEALYIF 401
Qy 359 ALTSYLYIYSNEYKMGPGGRNLYKIQLSDY-TKVTCLSCELNPERCQYYSVSFSKEAKY 417
Db 402 RVTQDSLFYSSNEFEFGYPGRRNIYRISIGNSPPSKCVTCHLRKERCQYYSFSYKAKY 461
Qy 418 YQURCSGPGPLPYTLTHSSVNDKGLRVLEDNSALDKMLQNVQMPSSKKLDFILNETKFWYQ 477
Db 462 YALICYGPGLPISLTHDGRDQEIQVLEENKELENALRNLIQLPVAVEIKKLEDGGMTFWYK 521
Qy 478 MILPPHFDKSKYPLLLDVTYAGPCSQKADTVFRLNWTATYLASTENIIVASFDCRGSGYQG 537
Db 522 MILPPQFDRSKYPLLIQVYGGPCSVSRVFAVNWISYLASKGWIADVGRGTAFQG 581
Qy 538 DKIMHAINRLGTVEVEDQIEAARQFSKMGFVGNKRIAIWGSYGGVYVSMVLGSGGVF 597
Db 582 DXFLHAVYRKLGVYEVEDQITAVRKFTMGFIDFKRIAIWGSYGGVYVSSALASGTGLF 641
Qy 598 KCGIAPVSRWEYSDVSYTYRYMGLPTPDNDLHYRNSVTMSRAENFKQVEYLLIHGTA 657
Db 642 KCGIAPVSSWEYASVYTERFMGLPTKODNLEHYKNSTVMARAEYFRNVVDYLLIHGTA 701
Qy 658 DDNVHFOQSAQISKALVDGVDFQAMWYTDDEHGIASSTAHOHIYTHMSHFIKQCFSL 715
Db 702 DDNVHFQNSAQIAKALVNAQVDFQAMWYSDQNHGISGRSQNHLYTHMTFLKQCFSL 759
```

Search completed: January 30, 2006, 14:20:09
Job time : 167 sec

THIS PAGE BLANK (USPTO)